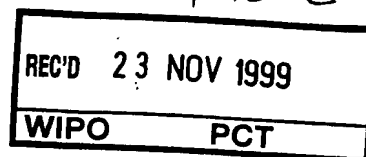


EP 99 / 7952

**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)



Bescheinigung

Die Haarmann & Reimer GmbH in Holzminden/Deutschland hat eine Patentanmeldung unter der Bezeichnung

"Konstruktion von Produktionsstämmen für die Herstellung von substituierten Phenolen durch gezielte Inaktivierung von Genen des Eugenol- und Ferulasäure-Katabolismus"

am 31. Oktober 1998 beim Deutschen Patent- und Markenamt eingereicht.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

Die Anmeldung hat im Deutschen Patent- und Markenamt vorläufig die Symbole C 12 N, C 12 P und C 07 H der Internationalen Patentklassifikation erhalten.

München, den 5. August 1999

Deutsches Patent- und Markenamt

Der Präsident

Im Auftrag

Waasmaier

Anzeichen: 198 50 242.7



Konstruktion von Produktionsstämmen für die Herstellung von substituierten Phenolen durch gezielte Inaktivierungen von Genen des Eugenol- und Ferulasäure-Katabolismus

5

Die vorliegende Erfindung betrifft die Konstruktion von Produktionsstämmen und ein Verfahren für die Herstellung substituierter Methoxyphenole, insbesondere Vanillin.

10

Die DE-A 4 227 076 (Verfahren zur Herstellung substituierter Methoxyphenole und dafür geeigneter Mikroorganismus) beschreibt die Herstellung substituierter Methoxyphenole mit einer neuen *Pseudomonas* sp.. Ausgangsmaterial ist hier Eugenol und die Produkte sind Ferulasäure, Vanillinsäure, Coniferylalkohol und Coniferylaldehyd.

15

Ebenfalls 1995 erscheint ein umfangreiches Review über die Biotransformationsmöglichkeiten mit Ferulasäure von Rosazza et al. (Biocatalytic transformation of ferulic acid: an abundant aromatic natural product; J. Ind. Microbiol. 15:457-471).

20

Die Gene und Enzyme zur Synthese von Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und Vanillinsäure aus *Pseudomonas* sp. wurden in EP-A 0 845 532 beschrieben.

25

Die Enzyme zur Umsetzung von *trans*-Ferulasäure zu *trans*-Feruloyl-SCoA Ester und weiter zum Vanillin, sowie das Gen für die Spaltung des Esters wurden vom Institute of Food Research, Norwich, GB, in WO 97/35999 beschrieben. 1998 erscheint der Inhalt des Patents auch als wissenschaftliche Publikationen (Gasson et al. 1998. Metabolism of ferulic acid to vanillin. J. Biol. Chem. 273:4163-4170; Narbad and Gasson 1998. Metabolism of ferulic acid via vanillin using a novel CoA-

30

dependent pathway in a newly isolated strain of *Pseudomonas fluorescens*. Microbiology 144:1397 - 1405).

Die DE-A 195 32 317 beschreibt die fermentative Gewinnung von Vanillin aus Ferulasäure mit *Amycolatopsis* sp. in hohen Ausbeuten.

- 5 Die bekannten Verfahren haben den Nachteil, daß entweder nur sehr geringe Ausbeuten an Vanillin erzielt werden, oder von relativ teuren Edukten ausgegangen wird. Bei dem letztgenannten Verfahren (DE-A 195 32 317) werden zwar hohe Ausbeuten erzielt, jedoch bedingt der Einsatz von *Pseudomonas* sp. HR199 und *Amycolatopsis*
-

- 10 sp. HR167 für die Biotransformation von Eugenol zu Vanillin eine zweistufige Fermentationsführung und somit einen erheblichen Kosten- und Zeitaufwand.

Aufgabe der vorliegenden Erfindung ist es daher, Organismen zu konstruieren, die in der Lage sind den preiswerten Rohstoff Eugenol in einem einstufigen Prozeß zu Vanillin umzusetzen.

- 15 Diese Aufgabe wird durch die Konstruktion von Produktionsstämmen ein- oder mehrzelliger Organismen gelöst, die dadurch gekennzeichnet sind, daß Enzyme des Eugenol- und/oder Ferulasäure-Katabolismus derart inaktiviert sind, daß eine Akkumulation der Intermediate Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure erfolgt.

- 20 Der Produktionsstamm kann einzellig oder mehrzellig sein. Demgemäß können Gegenstand der Erfindung Mikroorganismen, Pflanzen oder Tiere sein. Darüber hinaus können auch Extrakte die aus dem Produktionsstamm gewonnen werden zum Einsatz kommen. Erfindungsgemäß werden vorzugsweise einzellige Organismen eingesetzt. Hierbei kann es sich um Mikroorganismen, tierische oder pflanzliche Zellen handeln. Besonders bevorzugt ist erfindungsgemäß der Einsatz von Pilzen und Bakterien. Höchst bevorzugt sind Bakterienarten. Unter den Bakterien können insbesondere *Rhodococcus*-, *Pseudomonas*- und *Escherichia*-Arten nach Veränderung des
- 25
- 30 Eugenol- und/oder Ferulasäure-Katabolismus zum Einsatz kommen.

Die Gewinnung der erfindungsgemäß einsetzbaren Organismen kann im einfachsten Fall mittels bekannter, konventioneller mikrobiologischer Methoden erfolgen. So kann die Enzymaktivität der am Eugenol- und/oder Ferulasäure-Katabolismus beteiligten Proteine durch den Einsatz von Enzym-Hemmstoffen verändert werden.

5 Darüber hinaus kann die Enzymaktivität der am Eugenol- und/oder Ferulasäure-Katabolismus beteiligten Proteine durch Mutation der für diese Proteine kodierenden Gene verändert werden. Derartige Mutationen können nach klassischen Methoden ungerichtet erzeugt werden, wie beispielsweise durch UV-Bestrahlung oder mutationsauslösende Chemikalien.

10 Ebenso sind gentechnische Methoden zur Gewinnung der erfindungsgemäßen Organismen geeignet, wie Deletionen, Insertionen und/oder Nukleotid-Austausche. So können beispielsweise die Gene der Organismen mit Hilfe von anderen DNA-Elementen (Ω -Elemente) inaktiviert werden. Ebenso können mittels geeigneter Vektoren

15 Austausch der intakten Gene gegen veränderte und/oder inaktivierte Gen-Strukturen durchgeführt werden. Die zu inaktivierenden Gene und die für die Inaktivierung eingesetzten DNA-Elemente können dabei durch klassische Klonierungstechniken oder durch Polymerase-Kettenreaktionen (PCR) gewonnen werden.

20 In einer möglichen Ausgestaltung der Erfindung kann beispielsweise der Eugenol- sowie der Ferulasäure-Katabolismus durch Ω -Element-Insertion oder Einführen von Deletionen in entsprechende Gene verändert werden. Hierbei können die Funktionen der Gene, die für Dehydrogenasen, Synthetasen, Hydratasen-Aldolasen, Thiolasen oder Demethylasen kodieren, mittels der oben genannten gentechnischen Methoden

25 inaktiviert werden, so daß die Erzeugung der betreffenden Enzyme blockiert ist. Vorzugsweise handelt es sich um die Gene, die für Coniferylalkohol-Dehydrogenasen, Coniferylaldehyd-Dehydrogenasen, Ferulasäure-CoA-Synthetasen, Enoyl-CoA-Hydratasen-Aldolasen, beta-Ketothiolasen, Vanillin-Dehydrogenasen oder Vanillinsäure-Demethylasen kodieren. Ganz besonders bevorzugt sind Gene, die die in der

30 EP-A 0845532 angegebenen Aminosäuresequenzen kodieren und/oder deren Allelvariationen kodierenden Nukleotidsequenzen.

Gegenstand der Erfindung sind demgemäß auch Gen-Strukturen zur Herstellung transformierter Organismen und Mutanten.

5 Vorzugsweise werden Gen-Strukturen zur Gewinnung der Organismen und Mutanten eingesetzt, bei denen die für Dehydrogenasen, Synthetasen, Hydratasen-Aldolasen, Thiolasen oder Demethylasen kodierenden Nukleotidsequenzen inaktiviert sind. Besonders bevorzugt sind Gen-Strukturen, bei denen die für Coniferylalkohol-

10 Dehydrogenasen, Coniferylaldehyd-Dehydrogenasen, Ferulasäure-CoA-Synthetasen, Enoyl-CoA-Hydratasen-Aldolasen, beta-Ketothiolasen, Vanillin-Dehydrogenasen oder Vanillinsäure-Demethylasen kodierenden Nukleotidsequenzen inaktiviert sind. Ganz besonders bevorzugt sind Gen-Strukturen, die die in den Figuren 1a bis 1r angegebenen Strukturen mit den in den Figuren 2a bis 2r wiedergegebenen Nukleotidsequenzen und/oder deren Allelvariationen kodierenden Nukleotidsequenzen aufweisen. Besonders bevorzugt sind hierbei Nukleotidsequenzen von 1 bis 18.

20 Die Erfindung schließt auch die Teilsequenzen dieser Gen-Strukturen sowie funktionelle Äquivalente ein. Unter funktionellen Äquivalenten sind solche Derivate der DNA zu verstehen, bei denen einzelne Nukleobasen ausgetauscht worden sind (Wobbel austausche), ohne die Funktion zu ändern. Auch auf Proteinebene können Aminosäuren ausgetauscht werden, ohne daß eine Veränderung der Funktion die Folge ist.

25 Den Gen-Strukturen können ein oder mehrere DNA-Sequenzen vor- und/oder nachgeschaltet sein. Durch Klonierung der Gen-Strukturen sind Plasmide bzw. Vektoren erhältlich, die zur Transformation und/oder Transfektion eines Organismus und/oder zur konjugativen Übertragung in einen Organismus geeignet sind.

30 Gegenstand der Erfindung sind ferner Plasmide und/oder Vektoren zur Herstellung der erfindungsgemäßen transformierten Organismen und Mutanten. Diese enthalten demgemäß die beschriebenen Gen-Strukturen. Die vorliegende Erfindung betrifft

demgemäß auch Organismen, die die genannten Plasmide und/oder Vektoren enthalten.

Die Art der Plasmide und/oder Vektoren hängt von deren Einsatzzweck ab. Um z. B. die intakten Gene des Eugenol- und/oder Ferulasäure-Katabolismus in *Pseudomonaden* gegen die durch Omega-Elemente inaktivierten Gene austauschen zu können, benötigt man Vektoren, die einerseits in *Pseudomonaden* übertragen werden können (konjugativ übertragbare Plasmide), andererseits dort jedoch nicht repliziert werden

können und somit in *Pseudomonaden* instabil sind (sogenannte Suizid-Plasmide). DNA-Abschnitte, die mit Hilfe eines solchen Plasmidsystems in *Pseudomonaden* übertragen werden, können nur erhalten bleiben, wenn sie durch homologe Rekombination in das Genom der Bakterienzelle integriert werden.

Die beschriebenen Gen-Strukturen, Vektoren und Plasmide können zur Herstellung verschiedener transformierter Organismen oder Mutanten verwendet werden. Mittels der genannten Gen-Strukturen können intakte Nukleinsäuresequenzen gegen veränderte und/oder inaktivierte Gen-Strukturen ausgetauscht werden. In den durch Transformation oder Transfektion oder Konjugation erhältlichen Zellen erfolgt durch homologe Rekombination ein Austausch des intakten Gens gegen die veränderte und/oder inaktivierte Gen-Struktur, wodurch die resultierenden Zellen nur noch über die veränderte und/oder inaktivierte Gen-Struktur im Genom verfügen. So können erfindungsgemäß vorzugsweise Gene derart verändert und/oder inaktiviert werden, daß die betreffenden Organismen Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure zu erzeugen vermögen.

Erfindungsgemäß derart konstruierte Produktionsstämme sind beispielsweise Mutanten des Stammes *Pseudomonas* sp. HR199 (DSM 7063) der in der DE-A 4 227 076 und der EP-A 0845532 genau beschrieben wurde, wobei sich unter anderem die entsprechenden Genstrukturen aus den Figuren 1a bis 1r in Verbindung mit den Figuren 2a bis 2r ergeben:

1. *Pseudomonas* sp. HR199 $\text{caA}\Omega\text{Km}$, enthaltend das durch ΩKm inaktivierte *calA*-Gen an Stelle des intakten *calA*-Gens kodierend für Coniferylalkohol-Dehydrogenase (Fig. 1a; Fig. 2a).
2. *Pseudomonas* sp. HR199 $\text{calA}\Omega\text{Gm}$, enthaltend das durch ΩGm inaktivierte *calA*-Gen an Stelle des intakten *calA*-Gens kodierend für Coniferylalkohol-Dehydrogenase (Fig. 1b; Fig. 2b).
3. *Pseudomonas* sp. HR199 $\text{calA}\Delta$, enthaltend das durch Deletion inaktivierte *calA*-Gen an Stelle des intakten *calA*-Gens kodierend für Coniferylalkohol-Dehydrogenase (Fig. 1c; Fig. 2c).
4. *Pseudomonas* sp. HR199 $\text{calB}\Omega\text{Km}$, enthaltend das durch ΩKm inaktivierte *calB*-Gen an Stelle des intakten *calB*-Gens kodierend für Coniferylaldehyd-Dehydrogenase (Fig. 1d; Fig. 2d)
5. *Pseudomonas* sp. HR199 $\text{calB}\Omega\text{Gm}$, enthaltend das durch ΩGm inaktivierte *calB*-Gen an Stelle des intakten *calB*-Gens kodierend für Coniferylaldehyd-Dehydrogenase (Fig. 1e; Fig. 2e).
6. *Pseudomonas* sp. HR199 $\text{calB}\Delta$, enthaltend das durch Deletion inaktivierte *calB*-Gen an Stelle des intakten *calB*-Gens kodierend für Coniferylaldehyd-Dehydrogenase (Fig. 1f; Fig. 2f).
7. *Pseudomonas* sp. HR199 $\text{fcs}\Omega\text{Km}$, enthaltend das durch ΩKm inaktivierte *fcs*-Gen an Stelle des intakten *fcs*-Gens kodierend für Ferulasäure-CoA-Synthetase (Fig.1g; Fig. 2g).
8. *Pseudomonas* sp. HR199 $\text{fcs}\Omega\text{Gm}$, enthaltend das durch ΩGm inaktivierte *fcs*-Gen an Stelle des intakten *fcs*-Gens kodierend für Ferulasäure-CoA-Synthetase (Fig.1h; Fig. 2h).
9. *Pseudomonas* sp. HR199 $\text{fcs}\Delta$, enthaltend das durch Deletion inaktivierte *fcs*-Gen an Stelle des intakten *fcs*-Gens kodierend für Ferulasäure-CoA-Synthetase (Fig.1i; Fig. 2i).
10. *Pseudomonas* sp. HR199 $\text{ech}\Omega\text{Km}$, enthaltend das durch ΩKm inaktivierte *ech*-Gen an Stelle des intakten *ech*-Gens kodierend für Enoyl-CoA-Hydratase-Aldolase (Fig.1j; Fig. 2j).

11. *Pseudomonas* sp. HR199 $\text{ech}\Omega\text{Gm}$, enthaltend das durch ΩGm inaktivierte ech -Gen an Stelle des intakten ech -Gens kodierend für Enoyl-CoA-Hydratase-Aldolase (Fig. 1k; Fig. 2k).
12. *Pseudomonas* sp. HR199 $\text{ech}\Delta$, enthaltend das durch Deletion inaktivierte ech -Gen an Stelle des intakten ech -Gens kodierend für Enoyl-CoA-Hydratase-Aldolase (Fig. 1l; Fig. 2l).
13. *Pseudomonas* sp. HR199 $\text{aat}\Omega\text{Km}$, enthaltend das durch ΩKm inaktivierte aat -Gen an Stelle des intakten aat -Gens kodierend für beta-Ketothiolase (Fig. 1m; Fig. 2m).

14. *Pseudomonas* sp. HR199 $\text{aat}\Omega\text{Gm}$, enthaltend das durch ΩGm inaktivierte aat -Gen an Stelle des intakten aat -Gens kodierend für beta-Ketothiolase (Fig. 1n; Fig. 2n).
15. *Pseudomonas* sp. HR199 $\text{aat}\Delta$, enthaltend das durch Deletion inaktivierte aat -Gen an Stelle des intakten aat -Gens kodierend für beta-Ketothiolase (Fig. 1o; Fig. 2o).
16. *Pseudomonas* sp. HR199 $\text{vdh}\Omega\text{Km}$, enthaltend das durch ΩKm inaktivierte vdh -Gen an Stelle des intakten vdh -Gens kodierend für Vanillin-Dehydrogenase (Fig. 1p; Fig. 2p).
17. *Pseudomonas* sp. HR199 $\text{vdh}\Omega\text{Gm}$, enthaltend das durch ΩGm inaktivierte vdh -Gen an Stelle des intakten vdh -Gens kodierend für Vanillin-Dehydrogenase (Fig. 1q; Fig. 2q).
18. *Pseudomonas* sp. HR199 $\text{vdh}\Delta$, enthaltend das durch Deletion inaktivierte vdh -Gen an Stelle des intakten vdh -Gens kodierend für Vanillin-Dehydrogenase (Fig. 1r; Fig. 2r).
19. *Pseudomonas* sp. HR199 $\text{vdhB}\Omega\text{Km}$, enthaltend das durch ΩKm inaktivierte vdhB -Gen an Stelle des intakten vdhB -Gens kodierend für Vanillin-Dehydrogenase II.
20. *Pseudomonas* sp. HR199 $\text{vdhB}\Omega\text{Gm}$, enthaltend das durch ΩGm inaktivierte vdhB -Gen an Stelle des intakten vdhB -Gens kodierend für Vanillin-Dehydrogenase II.

21. *Pseudomonas* sp. HR199 $\nu dhB\Delta$, enthaltend das durch Deletion inaktivierte νdhB -Gen an Stelle des intakten νdhB -Gens kodierend für Vanillin-Dehydrogenase II.
22. *Pseudomonas* sp. HR199 $adh\Omega Km$, enthaltend das durch ΩKm inaktivierte adh -Gen an Stelle des intakten adh -Gens kodierend für Alkohol-Dehydrogenase.
23. *Pseudomonas* sp. HR199 $adh\Omega Gm$, enthaltend das durch ΩGm inaktivierte adh -Gen an Stelle des intakten adh -Gens kodierend für Alkohol-Dehydrogenase.
-
24. *Pseudomonas* sp. HR199 $adh\Delta$ enthaltend das durch Deletion inaktivierte adh -Gen an Stelle des intakten adh -Gens kodierend für Alkohol-Dehydrogenase.
25. *Pseudomonas* sp. HR199 $vanA\Omega Km$, enthaltend das durch ΩKm inaktivierte $vanA$ -Gen an Stelle des intakten $vanA$ -Gens kodierend für die α -Untereinheit der Vanillinsäure-Demethylase.
26. *Pseudomonas* sp. HR199 $vanA\Omega Gm$, enthaltend das durch ΩGm inaktivierte $vanA$ -Gen an Stelle des intakten $vanA$ -Gens kodierend für die α -Untereinheit der Vanillinsäure-Demethylase.
27. *Pseudomonas* sp. HR199 $vanA\Delta$, enthaltend das durch Deletion inaktivierte $vanA$ -Gen an Stelle des intakten $vanA$ -Gens kodierend für die α -Untereinheit der Vanillinsäure-Demethylase.
28. *Pseudomonas* sp. HR199 $vanB\Omega Km$, enthaltend das durch ΩKm inaktivierte $vanB$ -Gen an Stelle des intakten $vanB$ -Gens kodierend für die β -Untereinheit der Vanillinsäure-Demethylase.
29. *Pseudomonas* sp. HR199 $vanB\Omega Gm$, enthaltend das durch ΩGm inaktivierte $vanB$ -Gen an Stelle des intakten $vanB$ -Gens kodierend für die β -Untereinheit der Vanillinsäure-Demethylase.
30. *Pseudomonas* sp. HR199 $vanB\Delta$, enthaltend das durch Deletion inaktivierte $vanB$ -Gen an Stelle des intakten $vanB$ -Gens kodierend für die β -Untereinheit der Vanillinsäure-Demethylase.

Gegenstand der Erfindung ist außerdem ein Verfahren zur biotechnischen Herstellung von organischen Verbindungen. Insbesondere können mit diesem Verfahren Alkohole, Aldehyde und organische Säuren hergestellt werden. Vorzugsweise handelt es sich hierbei um Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und Vanillinsäure.

In dem erfindungsgemäßen Verfahren werden die oben beschriebenen Organismen eingesetzt. Zu den ganz besonders bevorzugten Organismen gehören Bakterien, insbesondere die *Pseudomonas*-Arten. Im einzelnen können die oben genannten *Pseudomonas*-Arten vorzugsweise für folgende Verfahren eingesetzt werden:

1. *Pseudomonas* sp. HR199calA Ω Km, *Pseudomonas* sp. HR199calA Ω Gm und *Pseudomonas* sp. HR199calA Δ zur Herstellung von Coniferylalkohol aus Eugenol.
2. *Pseudomonas* sp. HR199calB Ω Km, *Pseudomonas* sp. HR199calB Ω Gm und *Pseudomonas* sp. HR199calB Δ zur Herstellung von Coniferylaldehyd aus Eugenol oder Coniferylalkohol.
3. *Pseudomonas* sp. HR199fcs Ω Km, *Pseudomonas* sp. HR199fcs Ω Gm, *Pseudomonas* sp. HR199fcs Δ , *Pseudomonas* sp. HR199ech Ω Km, *Pseudomonas* sp. HR199ech Ω Gm und *Pseudomonas* sp. HR199ech Δ zur Herstellung von Ferulasäure aus Eugenol oder Coniferylalkohol oder Coniferylaldehyd.
4. *Pseudomonas* sp. HR199vdh Ω Km, *Pseudomonas* sp. HR199vdh Ω Gm, *Pseudomonas* sp. HR199vdhD, *Pseudomonas* sp. HR199vdh Ω GmvdhB Ω Km, *Pseudomonas* sp. HR199vdh Ω KmvdhB Ω Gm, *Pseudomonas* sp. HR199vdh Δ vdhB Ω Gm und *Pseudomonas* sp. HR199vdh Δ vdhB Ω Km zur Herstellung von Vanillin aus Eugenol oder Coniferylalkohol oder Coniferylaldehyd oder Ferulasäure.

5. *Pseudomonas* sp. HR199vanA Ω Km, *Pseudomonas* sp. HR199vanA Ω Gm, *Pseudomonas* sp. HR199vanA Δ , *Pseudomonas* sp. HR199vanB Ω Km, *Pseudomonas* sp. HR199vanB Ω Gm und *Pseudomonas* sp. HR199vanB Δ zur Herstellung von Vanillinsäure aus Eugenol oder Coniferylalkohol oder Coniferylaldehyd oder Ferulasäure oder Vanillin.

Bevorzugtes Substrat ist Eugenol. Jedoch kann der Zusatz weiterer Substrate oder sogar der Austausch des Eugenol gegen ein anderes Substrat möglich sein.

Als Nährmedium für die erfindungsgemäß eingesetzten Organismen kommen synthetische, halbsynthetische oder komplexe Kulturmedien in Betracht. Diese können kohlenstoffhaltige und stickstoffhaltige Verbindungen, anorganische Salze, gegebenenfalls Spurenelemente sowie Vitamine enthalten.

Als kohlenstoffhaltige Verbindungen können Kohlenhydrate, Kohlenwasserstoffe oder organische Grundchemikalien in Betracht kommen. Beispiele für vorzugsweise verwendbare Verbindungen sind Zucker, Alkohole bzw. Zuckeralkohole, organische Säuren oder komplexe Gemische.

Als Zucker kommt vorzugsweise Glucose in Betracht. Als organische Säuren können vorzugsweise Zitronensäure oder Essigsäure zum Einsatz kommen. Zu den komplexen Gemischen zählen z. B. Malzextrakt, Hefeextrakt, Casein oder Caseinhydrolysat.

Als stickstoffhaltige Substrate kommen anorganische Verbindungen in Betracht. Beispiele hierfür sind Nitrate und Ammoniumsalze. Ebenso können organische Stickstoffquellen zum Einsatz kommen. Hierzu zählen Hefeextrakt, Sojamehl, Casein, Caseinhydrolysat und Maisquellwasser.

Zu den einsetzbaren anorganischen Salzen zählen beispielsweise Sulfate, Nitrate, Chloride, Carbonate und Phosphate. Als Metalle enthalten die genannte Salze vorzugsweise Natrium, Kalium, Magnesium, Mangan, Calcium, Zink und Eisen.

- 5 Die Temperatur für die Kultivierung liegt vorzugsweise im Bereich von 5 bis 100°C. Besonders bevorzugt ist der Bereich von 15 bis 60°C, höchst bevorzugt sind 22 bis 37°C.

- 10 Der pH-Wert des Mediums beträgt bevorzugt 2 bis 12. Besonders bevorzugt ist der Bereich von 4 bis 8.

- 15 Grundsätzlich können für die Durchführung des erfindungsgemäßen Verfahrens alle dem Fachmann bekannten Bioreaktoren eingesetzt werden. Vorzugsweise kommen alle für Submersverfahren geeigneten Vorrichtungen in Betracht. Das heißt, es können erfindungsgemäß Gefäße ohne oder mit mechanischer Mischeinrichtung eingesetzt werden. Zu ersteren zählen z. B. Schüttelapparaturen, Blasensäulen- oder Schlaufenreaktoren. Zu letzteren gehören vorzugsweise alle bekannten Vorrichtungen mit Rührern in beliebiger Gestaltung.

- 20 Das erfindungsgemäße Verfahren kann kontinuierlich oder diskontinuierlich durchgeführt werden. Die Dauer der Fermentation bis zum Erreichen einer maximalen Produktmenge hängt von der speziellen Art des eingesetzten Organismus ab. Grundsätzlich liegen die Zeiten der Fermentation jedoch zwischen 2 und 200 Stunden.

- 25 Im folgenden wird die Erfindung unter Bezugnahme auf Beispiele näher erläutert:

- 30 Von dem Eugenol verwertenden Stamm *Pseudomonas* sp. HR199 (DSM 7063) wurden gezielt Mutanten erzeugt, wobei spezifisch Gene des Eugenol-Katabolismus durch Insertion von Omega-Elementen oder durch Einführen von Deletionen inaktiviert wurden. Als Omega-Elemente dienten DNA-Abschnitte die für Antibiotikaresistenzen gegen Kanamycin (Ω Km) und Gentamycin (Ω Gm) codierten. Diese

Resistenzgene wurden ausgehend von Tn5 und dem Plasmid pBBR1MCS-5 mit Hilfe von Standardmethoden isoliert. Die Gene *calA*, *calB*, *fcs*, *ech*, *aat*, *vdh*, *adh*, *vdhB*, *vanA* und *vanB*, die für Coniferylalkohol-Dehydrogenase, Coniferylaldehyd-Dehydrogenase, Ferulasäure-CoA Synthetase, Enoyl-CoA Hydratase-Aldolase, beta-Ketothiolase, Vanillin-Dehydrogenase, Alkohol-Dehydrogenase, Vanillin-Dehydrogenase II und Vanillinsäure-Demethylase codieren wurden ausgehend von genomischer DNA des Stammes *Pseudomonas* sp. HR199 mit Hilfe von Standardmethoden isoliert und in pBluescript SK⁻ kloniert. Aus diesen Genen wurden durch Verdauung

mit geeigneten Restriktionsendonukleasen DNA-Abschnitte entfernt (Deletion), bzw durch Ω -Elemente substituiert (Insertion), wodurch das jeweilige Gen inaktiviert wurde. Die auf diese Weise mutierten Gene wurden in konjugativ übertragbare Vektoren umklont und anschließend in den Stamm *Pseudomonas* sp. HR199 eingeführt. Durch geeignete Selektion wurden Transkonjuganten erhalten, die das jeweils funktionsfähige wildtyp-Gen gegen das neu eingebrachte inaktivierte Gen ausgetauscht hatten. Die so erhaltenen Insertions- und Deletionsmutanten wiesen nur noch das jeweils inaktivierte Gen auf. Auf diese Weise wurden sowohl Mutanten mit nur einem defekten Gen als auch Mehrfachmutanten, in denen mehrere Gene auf diese Weise inaktiviert wurden, erhalten. Diese Mutanten wurden für die Biotransformation von

- a) Eugenol zu Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure;
- b) Coniferylalkohol zu Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure;
- c) Coniferylaldehyd zu Ferulasäure, Vanillin und/oder Vanillinsäure;
- d) Ferulasäure zu Vanillin und/oder Vanillinsäure und
- e) Vanillin zu Vanillinsäureeingesetzt.

Material und Methoden

Wachstumsbedingungen der Bakterien.

5 Stämme von *Escherichia coli* wurden bei 37°C in Luria-Bertani (LB) oder M9-Mineralmedium (Sambrook, J., E. F. Fritsch und T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2. Aufl., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.) angezogen. Stämme von *Pseudomonas* sp. wurden bei 30°C in Nutrient Broth (NB, 0,8%, wt/vol) oder in Mineralmedium (MM) (Schlegel, H. G. et al. 1961 Arch. Mikrobiol. 38:209-222) bzw. HR-Mineralmedium (HR-MM)

10 (Rabenhorst, J. 1996. Appl. Microbiol. Biotechnol. 46:470-474.) angezogen. Ferulasäure, Vanillin, Vanillinsäure und Protocatechusäure wurden in Dimethylsulfoxid gelöst, und dem jeweiligen Medium in einer Endkonzentration von 0.1% (wt/vol) zugesetzt. Eugenol wurde dem Medium direkt in einer Endkonzentration von 0.1% (vol/vol) zugesetzt, bzw. in den Deckel von MM-Agarplatten auf Filterpapier

15 (Rundfilter 595, Schleicher & Schuell, Dassel, Deutschland) appliziert. Bei der Anzucht von Transkonjuganten und Mutanten von *Pseudomonas* sp. wurde Tetracyclin, Kanamycin, und Gentamycin in Endkonzentrationen von 25 µg/ml bzw. 100 µg/ml bzw. 7,5 µg/ml eingesetzt.

20 **Qualitativer und quantitativer Nachweis von Stoffwechselintermediaten in Kulturüberständen.**

Kulturüberstände wurden direkt, bzw. nach Verdünnung mit H₂O-bidest. mittels Hochdruck-Flüssigkeits-Chromatographie (Knauer-HPLC) analysiert. Die Chromatographie erfolgte an Nucleosil-100 C18 (7 µm, 250 x 4 mm). Als Lösungsmittel

25 diente 0.1% (vol/vol) Ameisensäure und Acetonitril. Der verwendete Gradient zur Elution der Substanzen verlief wie folgt:

00:00 - 06:30 → 26% Acetonitril

06:30 - 08:00 → 100% Acetonitril

30 08:00 - 12:00 → 100% Acetonitril

12:00 - 13:00 → 26% Acetonitril

13:00 - 18:00 → 26% Acetonitril

Reinigung der Vanillin-Dehydrogenase-II.

Die Aufreinigung erfolgte bei 4°C.

5 Rohextrakt.

Auf Eugenol angezogene Zellen von *Pseudomonas* sp. HR199 wurden in 10 mM Natriumphosphat-Puffer, pH 6.0 gewaschen, im gleichen Puffer resuspendiert und durch zweimalige Passage einer French-Presse (Amicon, Silver Spring, Maryland,

10

USA) bei einem Druck von 1000 psi aufgeschlossen. Das Zellhomogenat wurde einer Ultrazentrifugation (1 h, 100 000 x g, 4°C) unterzogen, wodurch die lösliche Fraktion des Rohextraktes als Überstand erhalten wurde.

Anionenaustauschchromatographie an DEAE-Sephacel.

15

Die lösliche Fraktion des Rohextraktes wurde über Nacht gegen 10 mM Natriumphosphat-Puffer, pH 6.0 dialysiert. Das Dialysat wurde auf eine mit 10 mM Natriumphosphat-Puffer, pH 6.0 äquilibrierte DEAE-Sephacel-Säule (2,6 cm x 35 cm, Bettvolumen[BV]: 186 ml) mit einer Durchflußrate von 0.8 ml/min aufgetragen. Die Säule wurde mit zwei BV 10 mM Natriumphosphat-Puffer, pH 6.0 gespült. Die Elution der Vanillin-Dehydrogenase-II (VDH-II) erfolgte mit einem linearen Salzgradient von 0 bis 400 mM NaCl in 10 mM Natriumphosphat-Puffer, pH 6.0 (750 ml). Es wurden 10 ml-Fractionen aufgefangen. Fractionen mit hoher VDH-II-Aktivität wurden zum DEAE-Pool vereinigt.

20

Bestimmung der Vanillin-Dehydrogenase-Aktivität.

25

Die Bestimmung der VDH-Aktivität erfolgte bei 30°C durch einen optisch enzymatischen Test. Der Reaktionsansatz mit einem Volumen von 1 ml enthielt 0.1 mmol Kalium-Phosphat (pH 7.1), 0.125 µmol Vanillin, 0.5 µmol NAD, 1.2 µmol Pyruvat (Na-Salz), Lactat-Dehydrogenase (1 U; aus Schweineherz) und Enzymlösung. Die Oxidation von Vanillin wurde bei $\lambda = 340$ nm verfolgt ($\epsilon_{\text{Vanillin}} = 11,6 \text{ cm}^2/\mu\text{mol}$).

30

Die Enzymaktivität wurde in Einheiten (U) angegeben, wobei 1 U der Enzymmenge entspricht, die 1 µmol Vanillin pro Minute umsetzt. Die Proteinkonzentrationen in

den Proben wurden nach Lowry et al. (Lowry, O. H., N. J. Rosebrough, A. L. Farr und R. J. Randall. 1951. J. Biol. Chem. **193**:265-275) bestimmt.

Bestimmung der Coniferylalkohol-Dehydrogenase-Aktivität.

5 Die Bestimmung der CADH-Aktivität erfolgte bei 30°C durch einen optisch enzymatischen Test nach Jaeger et al. (Jaeger, E., L. Eggeling und H. Sahn. 1981. Current Microbiology. **6**:333-336). Der Reaktionsansatz mit einem Volumen von 1 ml enthielt 0.2 mmol Tris/HCl (pH 9.0), 0.4 µmol Coniferylalkohol, 2 µmol NAD,

10 0.1 mmol Semicarbazid und Enzymlösung. Die Reduktion von NAD wurde bei $\lambda = 340$ nm verfolgt ($\epsilon = 6.3 \text{ cm}^2/\mu\text{mol}$). Die Enzymaktivität wurde in Einheiten (U) angegeben, wobei 1 U der Enzymmenge entspricht, die 1 µmol Substrat pro Minute umsetzt. Die Proteinkonzentrationen in den Proben wurden nach Lowry et al. (Lowry, O. H., N. J. Rosebrough, A. L. Farr und R. J. Randall. 1951. J. Biol. Chem. **193**:265-275) bestimmt.

15

Bestimmung der Coniferylaldehyd-Dehydrogenase-Aktivität.

Die Bestimmung der CALDH-Aktivität erfolgte bei 30°C durch einen optisch enzymatischen Test. Der Reaktionsansatz mit einem Volumen von 1 ml enthielt 0.1 mmol Tris/HCl (pH 8.8), 0.08 µmol Coniferylaldehyd, 2.7 µmol NAD und Enzymlösung.

20 Die Oxidation von Coniferylaldehyd zu Ferulasäure wurde bei $\lambda = 400$ nm verfolgt ($\epsilon = 34 \text{ cm}^2/\mu\text{mol}$). Die Enzymaktivität wurde in Einheiten (U) angegeben, wobei 1 U der Enzymmenge entspricht, die 1 µmol Substrat pro Minute umsetzt. Die Proteinkonzentrationen in den Proben wurden nach Lowry et al. (Lowry, O. H., N. J. Rosebrough, A. L. Farr und R. J. Randall. 1951. J. Biol. Chem. **193**:265-275) bestimmt.

25

Bestimmung der Ferulasäure-CoA-Synthetase (Ferulasäure-Thiokinase)-Aktivität.

30 Die Bestimmung der FCS-Aktivität erfolgte bei 30°C durch einen optisch enzymatischen Test, modifiziert nach Zenk et al. (Zenk et al. 1980. Anal. Biochem. **101**:182-187). Der Reaktionsansatz mit einem Volumen von 1 ml enthielt 0.09 mmol Kalium-

Phosphat (pH 7.0), 2.1 μmol MgCl_2 , 0.7 μmol Ferulasäure, 2 μmol ATP, 0.4 μmol Coenzym A und Enzymlösung. Die Entstehung des CoA-Esters aus Ferulasäure wurde bei $\lambda = 345 \text{ nm}$ verfolgt ($\epsilon = 10 \text{ cm}^2/\mu\text{mol}$). Die Enzymaktivität wurde in Einheiten (U) angegeben, wobei 1 U der Enzymmenge entspricht, die 1 μmol Substrat pro Minute umsetzt. Die Proteinkonzentrationen in den Proben wurden nach Lowry et al. (Lowry, O. H., N. J. Rosebrough, A. L. Farr und R. J. Randall. 1951. J. Biol. Chem. **193**:265-275) bestimmt.

Electrophoretische Methoden.

Die Auftrennung von proteinhaltigen Extrakten erfolgte in 7.4% (wt/vol) Polyacrylamidgelen unter nativen Bedingungen nach der Methode von Stegemann et al. (Stegemann et al. 1973. Z. Naturforsch. **28c**:722-732) und unter denaturierenden Bedingungen in 11.5% (wt/vol) Polyacrylamidgelen nach der Methode von Laemmli (Laemmli, U. K. 1970. Nature (London) **227**:680-685). Zur unspezifischen Proteinfärbung wurde Serva Blue R verwendet. Zur spezifischen Anfärbung der Coniferylalkohol-, Coniferylaldehyd- und Vanillin-Dehydrogenase wurden die Gele für 20 min in 100 mM KP-Puffer (pH 7.0) umgepuffert und anschließend bei 30°C im gleichen Puffer dem 0.08% (wt/vol) NAD, 0.04% (wt/vol) p-Nitroblau-Tetrazoliumchlorid, 0.003% (wt/vol) Phenazine-Methosulfat und 1 mM des jeweiligen Substrates zugesetzt worden war inkubiert, bis entsprechende Farbbanden sichtbar wurden.

Transfer von Proteinen aus Polyacrylamidgelen auf PVDF-Membranen. Proteine wurden aus SDS-Polyacrylamidgelen mit Hilfe eines Semidry-Fastblot Gerätes (B32/33, Biometra, Göttingen, Deutschland) nach Herstellerangaben auf PVDF-Membranen (Waters-Milipore, Bedford, Mass., USA) übertragen.

Bestimmung von N-terminalen Aminosäuresequenzen.

Die Bestimmung von N-terminalen Aminosäuresequenzen erfolgte mit Hilfe eines Protein Peptide Sequenzers (Typ 477 A, Applied Biosystems, Foster City, USA) und eines PTH-Analysers nach Herstellerangaben.

Isolierung und Manipulation von DNA.

Die Isolierung von genomischer DNA erfolgte nach der Methode von Marmur (Marmur, J. 1961. J. Mol. Biol. 3:208-218). Die Isolierung und Analyse von anderer Plasmid-DNA bzw. von DNA-Restriktionsfragmenten erfolgte nach Standardmethoden (Sambrook, J. E. F. Fritsch und T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2. Aufl., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.).

Transfer von DNA.

Die Präparation und Transformation von kompetenten *Escherichia coli*-Zellen erfolgte nach der Methode von Hanahan (Hanahan, D. 1983. J. Mol. Biol. 166:557-580). Konjugativer Plasmidtransfer zwischen Plasmid-tragenden *Escherichia coli* S17-1-Stämmen (Donor) und *Pseudomonas* sp.-Stämmen (Rezipient) erfolgte auf NB-Agarplatten nach der Methode von Friedrich et al. (Friedrich, B. et al. 1981. J. Bacteriol. 147:198-205), oder durch eine "Minikomplementations-Methode" auf MM-Agarplatten mit 0.5% (wt/vol) Gluconat als C-Quelle und 25 µg/ml Tetracyclin oder 100 µg/ml Kanamycin. Dabei wurden Zellen des Rezipienten in einer Richtung als Impfstrich aufgetragen. Nach 5 min wurden dann Zellen der Donor-Stämme als Impfstriche aufgetragen, wobei der Rezipienten-Impfstrich gekreuzt wurde. Nach einer Inkubation für 48 h bei 30°C wuchsen die Transkonjuganten direkt hinter der Kreuzungsstelle, wohingegen weder Donor- noch Rezipienten-Stamm zum Wachstum in der Lage war.

Hybridisierungsexperimente.

DNA-Restriktionsfragmente wurden in einem 0.8% (wt/vol) Agarose-Gel in 50 mM Tris- 50 mM Borsäure- 1.25 mM EDTA-Puffer (pH 8.5) elektrophoretisch aufgetrennt (Sambrook, J. E. F. Fritsch und T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2. Aufl., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.). Die Übertragung der denaturierten DNA aus dem Gel auf eine positiv geladene Nylonmembran (Porengröße: 0.45 µm, Pall Filtrationstechnik, Dreieich, Deutschland), die anschließende Hybridisierung mit biotinylierten, bzw.

Digoxigenin-markierten DNA-Sonden und die Herstellung dieser DNA-Sonden erfolgte nach Standardmethoden (Sambrook, J. E. F. Fritsch und T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2. Aufl., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.).

5

DNA-Sequenzierung.

Die Bestimmung von Nukleotidsequenzen erfolgte nach der Didesoxy-Kettenabbruch-Methode von Sanger et al. (Sanger et al. 1977. Proc. Natl. Acad. Sci. USA

74:5463-5467) "nicht-radioaktiv" mit einem "LI-COR DNA-Sequencer Modell 4000L" (LI-COR Inc., Biotechnology Division, Lincoln, NE, USA) unter Verwendung eines "Thermo Sequenase fluorescent labelled primer cycle sequencing kit with 7-deaza-dGTP" (Amersham Life Science, Amersham International pls, Little Chalfont, Buckinghamshire, England) jeweils nach Vorschrift des Herstellers.

15 Mit Hilfe von synthetischen Oligonukleotiden wurde nach der "Primer-hopping Strategie" von Strauss et al. (Strauss, E. C. et al. 1986. Anal. Biochem. 154:353-360) sequenziert.

Chemikalien, Biochemikalien und Enzyme.

20 Restriktionsenzyme, T4 DNA-Ligase, Lambda-DNA und Enzyme bzw. Substrate für die optisch enzymatischen Tests wurden von C. F. Boehringer & Söhne (Mannheim, Deutschland) oder von GIBCO/BRL (Eggenstein, Deutschland) bezogen. [γ -³²P]ATP kam von Amersham/Buchler (Braunschweig, Deutschland). Oligonukleotide wurden von der Firma MWG-Biotech GmbH (Ebersberg, Deutschland) bezogen.

25 Agarose vom Typ NA wurde von Pharmacia-LKB (Uppsala, Schweden) bezogen. Alle anderen Chemikalien waren von Haarmann & Reimer (Holzminden, Deutschland), E. Merck AG (Darmstadt, Deutschland), Fluka Chemie (Buchs., Schweiz), Serva Feinbiochemica (Heidelberg, Deutschland) oder Sigma Chemie (Deisenhofen, Deutschland).

Beispiele

Beispiel 1

5 **Konstruktion von Omega-Elementen, die Resistenzen gegenüber Kanamycin (Ω Km) bzw. Gentamycin (Ω Gm) vermitteln.**

Für die Konstruktion des Ω Km-Elements wurde das 2099 bp *Bgl*I-Fragment des Transposons Tn5 (Auerswald E., A. G. Ludwig und H. Schaller. 1981. Cold Spring

10 Harb. Symp. Quant. Biol. **45**:107-113; Beck E., G. Ludwig, E. A. Auerswald, B. Reiss und H. Schaller. 1982. Gene **19**:327-336; Mazodier P., P. Cossart, E. Giraud und F. Gasser. 1985. Nucleic Acids Res. **13**:195-205.) präparativ isoliert. Das Fragment wurde durch Behandlung mit der Nuklease Bal-31 auf ca. 990 bp verkürzt. Dieses Fragment, das nur noch das Kanamycin-Resistenzgen (codierend für eine Aminoglycosid-3'-O-Phosphotransferase) umfaßte, wurde anschließend mit *Sma*I geschnittener pSKsym-DNA (pBluescript SK⁻-Derivat, welches eine symetrisch aufgebaute multiple Klonierungsstelle [*Sal*I, *Hind*III, *Eco*RI, *Sma*I, *Eco*RI, *Hind*III, *Sal*I] enthält) ligiert. Aus dem resultierenden Plasmid konnte das Ω Km-Element als *Sma*I-, *Eco*RI-, *Hind*III- oder *Sal*I-Fragment reisoliert werden.

20 Für die Konstruktion des Ω Gm-Elements wurde das 983 bp *Eae*I-Fragment des Plasmids pBBR1MCS-5 (Kovach M.E., P. H. Elzer, D. S. Hill, G. T. Robertson, M. A. Farris, R. M. Roop und K. M. Peterson. 1995. Gene **166**:175-176.) präparativ isoliert und anschließend mit Mung Bean Nuklease (Abdauen von einzelsträngigen DNA-Molekülen) behandelt. Dieses Fragment, das nur noch das Gentamycin-Resistenzgen (codierend für eine Gentamycin-3-Acetyltransferase) umfaßte, wurde anschließend mit *Sma*I geschnittener pSKsym-DNA (s.o.) ligiert. Aus dem resultierenden Plasmid konnte das Ω Gm-Element als *Sma*I-, *Eco*RI-, *Hind*III- oder *Sal*I-Fragment reisoliert werden.

Beispiel 2

Klonierung der Gene aus *Pseudomonas* sp. HR199 (DSM7063), die durch Insertion von Ω -Elementen oder durch Deletionen inaktiviert werden sollten.

5 Die separaten Klonierungen der Gene *fcs*, *ech*, *vdh* und *aat* erfolgte ausgehend von den *E. coli* S17-1 Stämmen DSM 10439 und DSM 10440 mit den Plasmiden pE207 und pE5-1 (siehe EP-A 0845532). Aus diesen Plasmiden wurden die angegebenen Fragmente präparativ isoliert und wie im weiteren beschrieben behandelt:

10 Für die Klonierung des *fcs*-Gens wurde das 2350 bp große *SalI*/*EcoRI*-Fragment des Plasmids pE207 und das 3700 bp große *EcoRI*/*SalI*-Fragment des Plasmids pE5-1 zusammen in pBluescript SK⁻ in einer Weise kloniert, daß beide Fragmente über die *EcoRI*-Enden miteinander verbunden waren. Ausgehend von dem resultierenden Hybridplasmid wurde das 6050 bp *SalI*-Fragment präparativ isoliert und durch

15 Behandlung mit der Nuklease Bal-31 auf ca. 2480 bp verkürzt. Anschließend wurden an die Fragment-Enden *PstI*-Linker ligiert und das Fragment nach *PstI*-Verdauung in pBluescript SK⁻ kloniert (pSK_{*fcs*}). Nach Transformation von *E. coli* XL1-Blue wurden Klone erhalten, die das *fcs*-Gen exprimierten und eine FCS-Aktivität von 0.2 U/mg Protein aufwiesen.

20 Für die Klonierung des *ech*-Gens wurde das 3800 bp große *HindIII*/*EcoRI*-Fragment des Plasmids pE207 präparativ isoliert und durch Behandlung mit der Nuklease Bal-31 auf ca. 1470 bp verkürzt. Anschließend wurden an die Fragment-Enden *EcoRI*-Linker ligiert und das Fragment nach *EcoRI*-Verdauung in pBluescript SK⁻ kloniert

25 (pSK_{*ech*}).

Für die Klonierung des *vdh*-Gens wurde das 2350 bp große *SalI*/*EcoRI*-Fragment des Plasmids pE207 präparativ isoliert. Nach Klonierung in pBluescript SK⁻ wurde das

30 Fragment mit Hilfe eines Exonuklease III/ Mung Bean Nukleasesystems einseitig um ca. 1530 bp verkürzt. Anschließend wurde an das Fragmentende ein *EcoRI*-

Linker ligiert und das Fragment nach *EcoRI*-Verdauung in pBluescript SK⁻ kloniert (pSK*vdh*). Nach Transformation von *E. coli* XL1-Blue wurden Klone erhalten, die das *vdh*-Gen exprimierten und eine VDH-Aktivität von 0.01 U/mg Protein aufwiesen.

5

Für die Klonierung des *aat*-Gens wurde das 3700 bp große *EcoRI*/*SalI*-Fragment des Plasmids pE5-1 präparativ isoliert und durch Behandlung mit der Nuklease Bal-31 auf ca. 1590 bp verkürzt. Anschließend wurden an die Fragment-Enden *EcoRI*-

Linker ligiert und das Fragment nach *EcoRI*-Verdauung in pBluescript SK⁻ kloniert (pSK*aat*).

Beispiel 3

Inaktivierung der oben beschriebenen Gene durch Insertion von Ω -Elementen, bzw. durch Deletion von Teilbereichen dieser Gene.

15

Das Plasmid pSK*fcs*, welches das *fcs*-Gen enthielt wurde mit *Bss*HII verdaut, wodurch ein 1290 bp großes Fragment aus dem *fcs*-Gen herausgeschnitten wurde. Nach Religation wurde das Deletions-Derivat des *fcs*-Gens (*fcs* Δ) (siehe Abb. 1i und 2i) kloniert in pBluescript SK⁻ (pSK*fcs* Δ) erhalten. Darüber hinaus wurden nach Herausschneiden des Fragments die Omega-Elemente Ω Km und Ω Gm an dessen Stelle einligiert. Dadurch entstanden die Ω -inaktivierten Derivate des *fcs*-Gens (*fcs* Ω Km, siehe Abb. 1g und 2g) und (*fcs* Ω Gm, siehe Abb. 1h und 2h) kloniert in pBluescript SK⁻ (pSK*fcs* Ω Km und pSK*fcs* Ω Gm). In Rohextrakten der erhaltenen *E. coli* Klone, deren Hybridplasmide ein durch Deletion bzw. Ω -Element-Insertion inaktiviertes *fcs*-Gen aufwiesen, konnte keine FCS-Aktivität nachgewiesen werden.

20

25

Das Plasmid pSK*ech*, welches das *ech*-Gen enthielt, wurde mit *NruI* verdaut, wodurch ein 53 bp und ein 430 bp großes Fragment aus dem *ech*-Gen herausgeschnitten wurde. Nach Religation wurde das Deletions-Derivat des *ech*-Gens (*ech* Δ , siehe Abb. 1l und 2l) kloniert in pBluescript SK⁻ (pSK*ech* Δ) erhalten. Darüber

30

hinaus wurden nach Herausschneiden der Fragmente die Omega-Elemente ΩK_m und ΩG_m an deren Stelle einligiert. Dadurch entstanden die Ω -inaktivierten Derivate des *ech*-Gens (*ech* ΩK_m und *ech* ΩG_m) kloniert in pBluescript SK⁻ (pSK*ech* ΩK_m und pSK*ech* ΩG_m).

5

Das Plasmid pSK*vdh*, welches das *vdh*-Gen enthielt wurde mit *Bss*HII verdaut, wodurch ein 210 bp großes Fragment aus dem *vdh*-Gen herausgeschnitten wurde. Nach Religation wurde das Deletions-Derivat des *vdh*-Gens (*vdh* Δ , siehe Abb. 1o

10

und 2o) kloniert in pBluescript SK⁻ (pSK*vdh* Δ) erhalten. Darüber hinaus wurden nach Herausschneiden des Fragments die Omega-Elemente ΩK_m und ΩG_m an dessen Stelle einligiert. Dadurch entstanden die Ω -inaktivierten Derivate des *vdh*-Gens (*vdh* ΩK_m und *vdh* ΩG_m) kloniert in pBluescript SK⁻ (pSK*vdh* ΩK_m , siehe Abb. 1m und 2m) und (pSK*vdh* ΩG_m , siehe Abb. 1n und 2n). In Rohextrakten der erhaltenen *E. coli* Klone, deren Hybridplasmide ein durch Deletion bzw. Ω -Element-Insertion inaktiviertes *vdh*-Gen aufwiesen, konnte keine VDH-Aktivität nachgewiesen werden.

15

20

Das Plasmid pSK*aat*, welches das *aat*-Gen enthielt wurde mit *Bss*HII verdaut, wodurch ein 59 bp großes Fragment aus dem *aat*-Gen herausgeschnitten wurde. Nach Religation wurde das Deletions-Derivat des *aat*-Gens (*aat* Δ , siehe Abb. 1r und 2r) kloniert in pBluescript SK⁻ (pSK*aat* Δ) erhalten. Darüber hinaus wurden nach Herausschneiden des Fragments die Omega-Elemente ΩK_m und ΩG_m an dessen Stelle einligiert. Dadurch entstanden die Ω -inaktivierten Derivate des *aat*-Gens (*aat* ΩK_m , siehe Abb. 1p und 2p) und (*aat* ΩG_m , siehe Abb. 1q und 2q) kloniert in pBluescript SK⁻ (pSK*aat* ΩK_m und pSK*aat* ΩG_m).

25

Beispiel 4**Umklonieren der durch Ω -Elemente inaktivierten Gene in das konjugativ übertragbare "Suizid-Plasmid" pSUP202.**

5 Um die durch Ω -Elemente inaktivierten Gene in *Pseudomonas* sp. HR199 gegen die intakten Gene austauschen zu können, benötigt man einen Vektor, der einerseits in *Pseudomonaden* übertragen werden kann (konjugativ übertragbare Plasmide), andererseits dort jedoch nicht repliziert werden kann und somit in *Pseudomonaden* insta-

10 bil ist ("Suizid-Plasmid"). DNA-Abschnitte, die mit Hilfe eines solchen Plasmid-systems in *Pseudomonaden* übertragen werden, können nur erhalten bleiben, wenn sie durch homologe Rekombination (RecA-abhängige Rekombination) in das Genom der Bakterienzelle integriert werden. Im vorliegenden Fall wurde das "Suizid-Plasmid" pSUP202 (Simon et al. 1983. In: A. Pühler. Molecular genetics of the bacteria-plant interaction. Springer Verlag, Berlin, Heidelberg, New York, S. 98-15 106.) eingesetzt.

20 Die inaktivierten Gene *fcs* Ω Km und *fcs* Ω Gm wurden nach *Pst*I-Verdauung aus den Plasmiden pSK*fcs* Ω Km und pSK*fcs* Ω Gm isoliert und mit *Pst*I geschnittener pSUP202 DNA ligiert. Die Ligationsansätze wurden nach *E. coli* S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium mit Kanamycin bzw. Gentamycin. Es wurden Kanamycin-resistente Transformanden erhalten, deren Hybridplasmid (pSUP*fcs* Ω Km) das inaktivierte Gen *fcs* Ω Km enthielt. Das entsprechende Hybridplasmid (pSUP*fcs* Ω Gm) der Gentamycin- resistenten Transformanden enthielt das inaktivierte Gen *fcs* Ω Gm.

25 Die inaktivierten Gene *ech* Ω Km und *ech* Ω Gm wurden nach *Eco*RI-Verdauung aus den Plasmiden pSK*ech* Ω Km und pSK*ech* Ω Gm isoliert und mit *Eco*RI geschnittener pSUP202 DNA ligiert. Die Ligationsansätze wurden nach *E. coli* S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium mit Kanamycin 30 bzw. Gentamycin. Es wurden Kanamycin-resistente Transformanden erhalten, deren

Hybridplasmid (pSUPech Ω Km) das inaktivierte Gen *ech* Ω Km enthielt. Das entsprechende Hybridplasmid (pSUPech Ω Gm) der Gentamycin- resistenten Transformanden enthielt das inaktivierte Gen *ech* Ω Gm.

5 Die inaktivierten Gene *vdh* Ω Km und *vdh* Ω Gm wurden nach *Eco*RI-Verdauung aus den Plasmiden pSK*vdh* Ω Km und pSK*vdh* Ω Gm isoliert und mit *Eco*RI geschnittener pSUP202 DNA ligiert. Die Ligationsansätze wurden nach *E. coli* S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium mit Kanamycin

10 bzw. Gentamycin. Es wurden Kanamycin-resistente Transformanden erhalten, deren Hybridplasmid (pSUP*vdh* Ω Km) das inaktivierte Gen *vdh* Ω Km enthielt. Das entsprechende Hybridplasmid (pSUP*vdh* Ω Gm) der Gentamycin- resistenten Transformanden enthielt das inaktivierte Gen *vdh* Ω Gm.

15 Die inaktivierten Gene *aat* Ω Km und *aat* Ω Gm wurden nach *Eco*RI-Verdauung aus den Plasmiden pSK*aat* Ω Km und pSK*aat* Ω Gm isoliert und mit *Eco*RI geschnittener pSUP202 DNA ligiert. Die Ligationsansätze wurden nach *E. coli* S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium mit Kanamycin bzw. Gentamycin. Es wurden Kanamycin-resistente Transformanden erhalten, deren Hybridplasmid (pSUP*aat* Ω Km) das inaktivierte Gen *aat* Ω Km enthielt. Das entsprechende Hybridplasmid (pSUP*aat* Ω Gm) der Gentamycin- resistenten Transformanden enthielt das inaktivierte Gen *aat* Ω Gm.

Beispiel 5

25 **Umklonieren der durch Deletion inaktivierten Gene in das konjugativ übertragbare "Suizid-Plasmid" mit "sacB-Selektionssystem" pHE55.**

Um die durch Deletion inaktivierten Gene in *Pseudomonas* sp. HR199 gegen die intakten Gene austauschen zu können, benötigt man einen Vektor, der die schon für pSUP202 beschriebenen Eigenschaften aufweist. Da im Gegensatz zu den durch Ω -Element inaktivierten Genen bei durch Deletion inaktivierten Genen keine Selektion

tionsmöglichkeit (keine Antibiotika-Resistenz) für den erfolgten Austausch der Gene in *Pseudomonas* sp. HR199 besteht, mußte ein anderes Selektionssystem zur Anwendung kommen. Bei dem "*sacB*-Selektionssystem" wird das auszutauschende, durch Deletion inaktivierte Gen in einem Plasmid kloniert, welches neben einem Antibiotika-Resistenzgen auch über das *sacB*-Gen verfügt. Nach konjugativer Übertragung dieses Hybridplasmids in einen Pseudomonaden wird das Plasmid durch homologe Rekombination an der Stelle in das Genom integriert, an der sich das intakte Gen befindet (erster "Cross over"). Auf diese Weise entsteht ein "heterogener" Stamm, der sowohl über ein intaktes als auch über ein durch Deletion inaktiviertes Gen verfügt, welche durch die pHE55-DNA voneinander getrennt sind. Diese Stämme weisen die durch den Vektor codierte Resistenz auf und besitzen darüber hinaus ein aktives *sacB*-Gen. Durch ein zweites homologes Rekombinationsereignis (zweiter "Cross over"), soll nun die pHE55-DNA zusammen mit dem intakten Gen aus der genomischen DNA ausgegliedert werden. Durch dieses Rekombinationsereignis entsteht ein Stamm, der nur noch über das inaktivierte Gen verfügt. Darüber hinaus kommt es zum Verlust der pHE55-codierten Antibiotika-Resistenz und des *sacB*-Gens. Streicht man Stämme auf Saccharose-haltigen Medien aus, werden Stämme die das *sacB*-Gen exprimieren im Wachstum gehemmt, da das Genprodukt Saccharose zu einem Polymer umsetzt, welches im Periplasma der Zellen akkumuliert wird. Zellen, die durch das zweite Rekombinationsereignis das *sacB*-Gen nicht mehr tragen, werden somit nicht im Wachstum gehemmt. Um eine phänotypische Selektionsmöglichkeit auf die Integration des durch Deletion inaktivierten Gens zu besitzen, tauscht man dieses nicht gegen ein intaktes Gen aus, sondern man bedient sich eines Stammes, in dem das auszutauschende Gen bereits durch Insertion eines Ω -Elements "markiert" vorliegt. Bei erfolgreichem Austausch verliert der resultierende Stamm die durch das Ω -Element codierte Antibiotika-Resistenz.

Das inaktivierte Gen *fcs* Δ wurden nach *Pst*I-Verdauung aus dem Plasmid pSK*fcs* Δ isoliert und mit *Pst*I geschnittener pHE55 DNA ligiert. Der Ligationsansatz wurde nach *E. coli* S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-

Medium. Es wurden Tetracyclin-resistente Transformanden erhalten, deren Hybridplasmid (pHE*fcs*Δ) das inaktivierte Gen *fcs*Δ enthielt.

5 Das inaktivierte Gen *ech*Δ wurden nach *Eco*RI-Verdauung aus dem Plasmid pSK*ech*Δ isoliert und mit Mung Bean Nuklease behandelt (Erzeugung von glatten Enden ["blunt ends"]). Das Fragment wurde mit *Bam*HI geschnittener und Mung Bean Nuklease behandelter pHE55 DNA ligiert. Der Ligationsansatz wurde nach *E. coli* S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-

10 Medium. Es wurden Tetracyclin-resistente Transformanden erhalten, deren Hybridplasmid (pHE*ech*Δ) das inaktivierte Gen *ech*Δ enthielt.

Das inaktivierte Gen *vdh*Δ wurden nach *Eco*RI-Verdauung aus dem Plasmid pSK*vdh*Δ isoliert und mit Mung Bean Nuklease behandelt. Das Fragment wurde mit *Bam*HI geschnittener und Mung Bean Nuklease behandelter pHE55 DNA ligiert. Der
15 Ligationsansatz wurde nach *E. coli* S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium. Es wurden Tetracyclin-resistente Transformanden erhalten, deren Hybridplasmid (pHE*vdh*Δ) das inaktivierte Gen *vdh*Δ enthielt.

Das inaktivierte Gen *aat*Δ wurden nach *Eco*RI-Verdauung aus dem Plasmid pSK*aat*Δ isoliert und mit Mung Bean Nuklease behandelt. Das Fragment wurde mit *Bam*HI geschnittener und Mung Bean Nuklease behandelter pHE55 DNA ligiert. Der
20 Ligationsansatz wurde nach *E. coli* S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium. Es wurden Tetracyclin-resistente Transformanden erhalten, deren Hybridplasmid (pHE*aat*Δ) das inaktivierte Gen *aat*Δ enthielt.

25

Beispiel 6

Erzeugung von Mutanten des Stammes *Pseudomonas* sp. HR199, bei denen spezifisch Gene des Eugenol-Katabolismus durch Insertion eines Ω -Elementes inaktiviert wurden.

Der Stamm *Pseudomonas* sp. HR199 wurde als Rezipient in Konjugationsexperimenten eingesetzt, bei denen Stämme von *E. coli* S17-1 als Donoren eingesetzt wurden, die die unten aufgeführten Hybridplasmide von pSUP202 enthielten. Die

Transkonjuganten wurden auf Gluconat-haltigem Mineralmedium selektiert, welches das dem Ω -Element entsprechende Antibiotikum enthielt. "Homogenote" (Austausch des intakten Gens gegen das durch Ω -Element-Insertion inaktivierte Gen durch doppeltes "Cross over") und "heterogenote" (Integration des Hybridplasmids in das Genom durch einfachen "Cross over") Transkonjuganten konnten anhand der durch pSUP202 codierten Tetracyclin-Resistenz unterschieden werden.

Die Mutanten *Pseudomonas* sp. HR199 *fcs* Ω Km und *Pseudomonas* sp. HR199 *fcs* Ω Gm wurden nach Konjugation von *Pseudomonas* sp. HR199 mit *E. coli* S17-1 (pSUP*fcs* Ω Km) bzw. *E. coli* S17-1 (pSUP*fcs* Ω Gm) erhalten. Der Austausch des intakten *fcs*-Gens gegen das durch Ω Km bzw. Ω Gm inaktivierte Gen (*fcs* Ω Km bzw. *fcs* Ω Gm) wurde mittels DNA-Sequenzierung verifiziert.

Die Mutanten *Pseudomonas* sp. HR199 *ech* Ω Km und *Pseudomonas* sp. HR199 *ech* Ω Gm wurden nach Konjugation von *Pseudomonas* sp. HR199 mit *E. coli* S17-1 (pSUP*ech* Ω Km) bzw. *E. coli* S17-1 (pSUP*ech* Ω Gm) erhalten. Der Austausch des intakten *ech*-Gens gegen das durch Ω Km bzw. Ω Gm inaktivierte Gen (*ech* Ω Km bzw. *ech* Ω Gm) wurde mittels DNA-Sequenzierung verifiziert.

Die Mutanten *Pseudomonas* sp. HR199 *vdh* Ω Km und *Pseudomonas* sp. HR199 *vdh* Ω Gm wurden nach Konjugation von *Pseudomonas* sp. HR199 mit *E. coli* S17-1 (pSUP*vdh* Ω Km) bzw. *E. coli* S17-1 (pSUP*vdh* Ω Gm) erhalten. Der Austausch des

intakten *vdh*-Gens gegen das durch Ω Km bzw. Ω Gm inaktivierte Gen (*vdh* Ω Km bzw. *vdh* Ω Gm) wurde mittels DNA-Sequenzierung verifiziert.

- 5 Die Mutanten *Pseudomonas* sp. HR199 *aat* Ω Km und *Pseudomonas* sp. HR199 *aat* Ω Gm wurden nach Konjugation von *Pseudomonas* sp. HR199 mit *E. coli* S17-1 (pSUP*aat* Ω Km) bzw. *E. coli* S17-1 (pSUP*aat* Ω Gm) erhalten. Der Austausch des intakten *aat*-Gens gegen das durch Ω Km bzw. Ω Gm inaktivierte Gen (*aat* Ω Km bzw. *aat* Ω Gm) wurde mittels DNA-Sequenzierung verifiziert.
-

- 10 Die Mutante *Pseudomonas* sp. HR199 *fcs* Ω Km*vdh* Ω Gm wurden nach Konjugation von *Pseudomonas* sp. HR199 *fcs* Ω Km mit *E. coli* S17-1 (pSUP*vdh* Ω Gm) erhalten. Der Austausch des intakten *vdh*-Gens gegen das durch Ω Gm inaktivierte Gen (*vdh* Ω Gm) wurde mittels DNA-Sequenzierung verifiziert.

- 15 Die Mutante *Pseudomonas* sp. HR199 *vdh* Ω Km*aat* Ω Gm wurden nach Konjugation von *Pseudomonas* sp. HR199 *vdh* Ω Km mit *E. coli* S17-1 (pSUP*aat* Ω Gm) erhalten. Der Austausch des intakten *aat*-Gens gegen das durch Ω Gm inaktivierte Gen (*aat* Ω Gm) wurde mittels DNA-Sequenzierung verifiziert.

- 20 Die Mutante *Pseudomonas* sp. HR199 *vdh* Ω Km*ech* Ω Gm wurden nach Konjugation von *Pseudomonas* sp. HR199 *vdh* Ω Km mit *E. coli* S17-1 (pSUP*ech* Ω Gm) erhalten. Der Austausch des intakten *ech*-Gens gegen das durch Ω Gm inaktivierte Gen (*ech* Ω Gm) wurde mittels DNA-Sequenzierung verifiziert.

Beispiel 7

Erzeugung von Mutanten des Stammes *Pseudomonas* sp. HR199, bei denen spezifisch Gene des Eugenol-Katabolismus durch Deletion eines Teilbereiches inaktiviert wurden.

Die Stämme *Pseudomonas* sp. HR199 *fcs*ΩKm, *Pseudomonas* sp. HR199 *ech*ΩKm, *Pseudomonas* sp. HR199 *vdh*ΩKm und *Pseudomonas* sp. HR199 *aat*ΩKm wurden als Rezipient in Konjugationsexperimenten eingesetzt, bei denen Stämme von *E. coli*

S17-1 als Donoren eingesetzt wurden, die die unten aufgeführten Hybridplasmide von pHE55 enthielten. Die "heterogenoten" Transkonjuganten wurden auf Gluconat-haltigem Mineralmedium selektiert, welches neben Tetracyclin (pHE55 codierte Resistenz) das dem Ω-Element entsprechende Antibiotikum enthält. Nach Ausstreichen auf Saccharose-haltigem Mineralmedium wurden Transkonjuganten erhalten, die durch ein zweites Rekombinationsereignis (zweiter "Cross over") die Vektor-DNA eliminiert hatten. Durch Ausstreichen auf Mineralmedium ohne Antibiotika bzw. mit dem Ω-Element entsprechenden Antibiotikum konnten die Mutanten erkannt werden, bei denen das durch Ω-Element inaktivierte Gen gegen das durch Deletion inaktivierte Gen ausgetauscht worden war (keine Antibiotika-Resistenz).

Die Mutante *Pseudomonas* sp. HR199 *fcs*Δ wurde nach Konjugation von *Pseudomonas* sp. HR199 *fcs*ΩKm mit *E. coli* S17-1 (pHE*fcs*Δ) erhalten. Der Austausch des durch ΩKm inaktivierten Gens (*fcs*ΩKm) gegen das durch Deletion inaktivierte Gen (*fcs*Δ) wurde mittels DNA-Sequenzierung verifiziert.

Die Mutanten *Pseudomonas* sp. HR199 *ech*Δ wurde nach Konjugation von *Pseudomonas* sp. HR199 *ech*ΩKm mit *E. coli* S17-1 (pHE*ech*Δ) erhalten. Der Austausch des durch ΩKm inaktivierten Gens (*ech*ΩKm) gegen das durch Deletion inaktivierte Gen (*ech*Δ) wurde mittels DNA-Sequenzierung verifiziert.

Die Mutanten *Pseudomonas* sp. HR199 *vdh* Δ wurde nach Konjugation von *Pseudomonas* sp. HR199 *vdh* Ω Km mit *E. coli* S17-1 (pHE*vdh* Δ) erhalten. Der Austausch des durch Ω Km inaktivierten Gens (*vdh* Ω Km) gegen das durch Deletion inaktivierte Gen (*vdh* Δ) wurde mittels DNA-Sequenzierung verifiziert.

5

Die Mutanten *Pseudomonas* sp. HR199 *aat* Δ wurde nach Konjugation von *Pseudomonas* sp. HR199 *aat* Ω Km mit *E. coli* S17-1 (pHE*aat* Δ) erhalten. Der Austausch des durch Ω Km inaktivierten Gens (*aat* Ω Km) gegen das durch Deletion inaktivierte Gen (*aat* Δ) wurde mittels DNA-Sequenzierung verifiziert.

10

Beispiel 8

Biotransformation von Eugenol zu Vanillin mit der Mutante *Pseudomonas* sp. HR199 *vdh* Ω Km.

15

Der Stamm *Pseudomonas* sp. HR199 *vdh* Ω Km wurde in 50 ml HR-MM mit 6 mM Eugenol bis zu einer optischen Dichte von ca. OD_{600nm} = 0.6 angezogen. Nach 17 h waren 2.9 mM Vanillin, 1.4 mM Ferulasäure und 0.4 mM Vanillinsäure im Kulturüberstand nachweisbar.

20

Beispiel 9

Biotransformation von Eugenol zu Ferulasäure mit der Mutante *Pseudomonas* sp. HR199 *vdh* Ω G*maat* Ω Km.

25

Der Stamm *Pseudomonas* sp. HR199 *vdh* Ω G*maat* Ω Km wurde in 50 ml HR-MM mit 6 mM Eugenol bis zu einer optischen Dichte von ca. OD_{600nm} = 0.6 angezogen. Nach 18 h waren 1.9 mM Vanillin, 2.4 mM Ferulasäure und 0.6 mM Vanillinsäure im Kulturüberstand nachweisbar.

Beispiel 10**Biotransformation von Eugenol zu Coniferylalkohol mit der Mutante *Pseudomonas* sp. HR199 *vdh*Ω*Gmaat*Ω*Km*.**

- 5 Der Stamm *Pseudomonas* sp. HR199 *vdh*Ω*Gmaat*Ω*Km* wurde in 50 ml HR-MM mit 6 mM Eugenol bis zu einer optischen Dichte von ca. OD_{600nm} = 0.4 angezogen. Nach 15 h waren 1.7 mM Coniferylalkohol, 1.4 mM Vanillin, 1.4 mM Ferulasäure und 0.2 mM Vanillinsäure im Kulturüberstand nachweisbar.
-

10

Beispiel 11**Fermentative Produktion von natürlichem Vanillin aus Eugenol im 10 l Fermenter mit der Mutante *Pseudomonas* sp. HR 199 *vdh*Ω*KM*.**

- 15 Mit 100 ml einer 24 Stunden alten Vorkultur, die auf einer Schüttelmaschine (120 Upm) bei 32°C in einem auf pH 7,0 eingestellten Medium aus 12,5 g/l Glycerin, 10 g/l Hefeextrakt und 0,37 g/l Essigsäure angezogen wurde, wurde der Produktionsfermenter beimpft. Der Fermenter enthielt 9,9 l Medium mit folgender Zusammensetzung: 1,5 g/l Hefeextrakt, 1,6 g/l KH₂PO₄, 0,2 g/l NaCl, 0,2 g/l MgSO₄. Der pH-Wert wurde mit Natronlauge auf pH 7,0 eingestellt. Nach der Sterilisation wurde
- 20 dem Medium 4 g Eugenol zugefügt. Die Temperatur betrug 32°C, die Belüftung 3 NI/min und die Rührerdrehzahl 600 Upm. Der pH-Wert wurde mit Natronlauge bei pH 6,5 gehalten.

- 25 Vier Stunden nach dem Animpfen wurde mit der kontinuierlichen Zugabe von Eugenol begonnen, so daß am Ende der Fermentation nach 65 Stunden 255 g Eugenol zur Kultur gegeben worden waren. Außerdem wurden während der Fermentation 40 g Hefeextrakt zugefüttert. Die Konzentration an Eugenol lag am Ende der Fermentation bei 0,2 g/l. Der Gehalt an Vanillin betrug 2,6 g/l. Zusätzlich lagen noch
- 30 3,4 g/l Ferulasäure vor.

Das so erhaltene Vanillin kann durch bekannte physikalische Verfahren wie Chromatographie, Destillation und/oder Extraktion isoliert und zur Herstellung natürlicher Aromen verwendet werden.

5 Erläuterungen zu den Figuren:

FIG. 1a bis 1r:

Gen- Strukturen zur Gewinnung von Organismen und Mutanten

10

*calA**: Teil des inaktivierten Gens der Coniferylalkohol-Dehydrogenase

*calB**: Teil des inaktivierten Gens der Coniferylaldehyd-Dehydrogenase

*fcs**: Teil des inaktivierten Gens der Ferulasäure-CoA Synthetase

*ech**: Teil des inaktivierten Gens der Enoyl-CoA Hydratase-Aldolase

15

*vdh**: Teil des inaktivierten Gens der Vanillin-Dehydrogenase

*aat**: Teil des inaktivierten Gens der beta-Ketothiolase

Die mit "*" versehenen Restriktionsenzym-Schnittstellen kamen für die Konstruktion zum Einsatz, sind jedoch in dem resultierenden Konstrukt nicht mehr funktionsfähig.

20

FIG. 2a: Nukleotidsequenz der Gen-Struktur *calA*ΩKm

FIG. 2b: Nukleotidsequenz der Gen-Struktur *calA*ΩGm:

FIG. 2c: Nukleotidsequenz der Gen-Struktur *calA*Δ

FIG. 2d: Nukleotidsequenz der Gen-Struktur *calB*ΩKm

5 FIG. 2e: Nukleotidsequenz der Gen-Struktur *calB*ΩGm

FIG. 2f: Nukleotidsequenz der Gen-Struktur *calB*Δ

FIG. 2g: Nukleotidsequenz der Gen-Struktur *fcs*ΩKm

FIG. 2h: Nukleotidsequenz der Gen-Struktur *fcs*ΩGm

FIG. 2i: Nukleotidsequenz der Gen-Struktur *fcs*Δ

10 FIG. 2j: Nukleotidsequenz der Gen-Struktur *ech*ΩKm

FIG. 2k: Nukleotidsequenz der Gen-Struktur *ech*ΩGm

FIG. 2l: Nukleotidsequenz der Gen-Struktur *ech*Δ

FIG. 2m: Nukleotidsequenz der Gen-Struktur *vdh*ΩKm

FIG. 2n: Nukleotidsequenz der Gen-Struktur *vdh*ΩGm

15 FIG. 2o: Nukleotidsequenz der Gen-Struktur *vdh*Δ

FIG. 2p: Nukleotidsequenz der Gen-Struktur *aat*ΩKm

FIG. 2q: Nukleotidsequenz der Gen-Struktur *aat*ΩGm

FIG. 2r: Nukleotidsequenz der Gen-Struktur *aat*Δ

Sequenzen

CTGCAGCCAG GGCTGAAAAG GAGGGATTCA GTGAGGTCAT GAAGGGAGGG GACGGCGCCT 60
 GGCTCCAATT GCTCGATGGC GCCGCGATTG AGTGTCTTGG GCGCGGTCTT GGAGAGTTCTG 120
 GCTAGGGAGA TAAATTTGCT GGCCATGGTG GCGGCCCTG ATGGGTTGGA TGATTTTCTG 180
 CATTCTGCAT CATGAAATTC ATGAAATCAT CACTTTTCGG GGGGTGGGTG CACGGGATTG 240
 AAGGTTGCTA GGAGAGTGCA TTGCTCGTAA GCCCAGGAAG CACGCGGGTT TCAGGATGGT 300
 GCATGGAAAT GGCATGAGCT TTGCTGGATA TGATTAGAGA CATTAACTAT TTTGGCGGAA 360
 TGAAGCACG ATTCCTCGCC CGGTAGAGCG GTAACCGCGA CATTCAAGAC CGTAAAAAGG 420

AAAGAGCATG CAA CTG ACC AAC AAG AAA ATC GTC GTC ACC GGA GTG TCC TCC 472
 Met Gln Leu Thr Asn Lys Lys Ile Val Val Thr Gly Val Ser Ser
 1 5 10 15

GT ATC GGT GCC GAA ACT GCC CGC GTT CTG CGC TCT CAC GGC GCC ACA 520
 Gly Ile Gly Ala Glu Thr Ala Arg Val Leu Arg Ser His Gly Ala Thr
 20 25 30

GTG ATT GGC GTA GAT CGC AAC ATG CCG AGC CTG ACT CTG GAT GCT TTC 568
 Val Ile Gly Val Asp Arg Asn Met Pro Ser Leu Thr Leu Asp Ala Phe
 35 40 45

GTT CAG GCT GAC CTG AGC CAT CCT GAA GGC ATC GAT AAG GCC ATC GGG 616
 Val Gln Ala Asp Leu Ser His Pro Glu Gly Ile Asp Lys Ala Ile
 50 55 60 62

ACAGCAAGCG AACCGAATT GCCAGCTGGG GCGCCCTCTG GTAAGGTTGG GAAGCCCTGC 676

AAAGTAAACT GGATGGCTTT CTGCGGCCA AGGATCTGAT GGCGCAGGGG ATCAAGATCT 736

GATCAAGAGA CAGGATGAGG ATCGTTTCGC ATG ATT GAA CAA GAT GGA TTG CAC 790
 Met Ile Glu Gln Asp Gly Leu His
 1 5

GCA GGT TCT CCG GCC GCT TGG GTG GAG AGG CTA TTC GGC TAT GAC TGG 838
 Ala Gly Ser Pro Ala Ala Trp Val Glu Arg Leu Phe Gly Tyr Asp Trp
 10 15 20

GCA CAA CAG ACA ATC GGC TGC TCT GAT GCC GCC GTG TTC CGG CTG TCA 886
 Ala Gln Gln Thr Ile Gly Cys Ser Asp Ala Ala Val Phe Arg Leu Ser
 25 30 35 40

GCG CAG GGG CGC CCG GTT CTT TTT GTC AAG ACC GAC CTG TCC GGT GCC 934
 Ala Gln Gly Arg Pro Val Leu Phe Val Lys Thr Asp Leu Ser Gly Ala
 45 50 55

CTG AAT GAA CTG CAG GAC GAG GCA GCG CGG CTA TCG TGG CTG GCC ACG 982
 Leu Asn Glu Leu Gln Asp Glu Ala Ala Arg Leu Ser Trp Leu Ala Thr
 60 65 70

ACG GGC GTT CCT TGC GCA GCT GTG CTC GAC GTT GTC ACT GAA GCG GGA	1030
Thr Gly Val Pro Cys Ala Ala Val Leu Asp Val Val Thr Glu Ala Gly	
75 80 85	
AGG GAC TGG CTG CTA TTG GGC GAA GTG CCG GGG CAG GAT CTC CTG TCA	1078
Arg Asp Trp Leu Leu Leu Gly Glu Val Pro Gly Gln Asp Leu Leu Ser	
90 95 100	
TCT CAC CTT GCT CCT GCC GAG AAA GTA TCC ATC ATG GCT GAT GCA ATG	1126
Ser His Leu Ala Pro Ala Glu Lys Val Ser Ile Met Ala Asp Ala Met	
105 110 115 120	
CGG CGG CTG CAT ACG CTT GAT CCG GCT ACC TGC CCA TTC GAC CAC CAA	1174
Arg Arg Leu His Thr Leu Asp Pro Ala Thr Cys Pro Phe Asp His Gln	
125 130 135	
GCG AAA CAT CGC ATC GAG CGA GCA CGT ACT CGG ATG GAA GCC GGT CTT	1222
Ala Lys His Arg Ile Glu Arg Ala Arg Thr Arg Met Glu Ala Gly Leu	
140 145 150	
GTC GAT CAG GAT GAT CTG GAC GAA GAG CAT CAG GGG CTC GCG CCA GCC	1270
Val Asp Gln Asp Asp Leu Asp Glu Glu His Gln Gly Leu Ala Pro Ala	
155 160 165	
GAA CTG TTC GCC AGG CTC AAG GCG CGC ATG CCC GAC GGC GAG GAT CTC	1318
Glu Leu Phe Ala Arg Leu Lys Ala Arg Met Pro Asp Gly Glu Asp Leu	
170 175 180	
GTC GTG ACC CAT GGC GAT GCC TGC TTG CCG AAT ATC ATG GTG GAA AAT	1366
Val Val Thr His Gly Asp Ala Cys Leu Pro Asn Ile Met Val Glu Asn	
185 190 195 200	
GGC CGC TTT TCT GGA TTC ATC GAC TGT GGC CGG CTG GGT GTG GCG GAC	1414
Gly Arg Phe Ser Gly Phe Ile Asp Cys Gly Arg Leu Gly Val Ala Asp	
205 210 215	
CGC TAT CAG GAC ATA GCG TTG GCT ACC CGT GAT ATT GCT GAA GAG CTT	1462
Arg Tyr Gln Asp Ile Ala Leu Ala Thr Arg Asp Ile Ala Glu Glu Leu	
220 225 230	
GGC GGC GAA TGG GCT GAC CGC TTC CTC GTG CTT TAC GGT ATC GCC GCT	1510
Gly Gly Glu Trp Ala Asp Arg Phe Leu Val Leu Tyr Gly Ile Ala Ala	
235 240 245	
CCC GAT TCG CAG CGC ATC GCC TTC TAT CGC CTT CTT GAC GAG TTC TTC	1558
Pro Asp Ser Gln Arg Ile Ala Phe Tyr Arg Leu Leu Asp Glu Phe Phe	
250 255 260 264	
TGAGCGGGAC TCTGGGGTTC GAAATGACCG ACCAAGCGAC GCCCTG GCC GCG GTG	1613
Ala Ala Val	
225	
ATT GCA TTC ATG TGT GCT GAG GAG TCA CGT TGG ATC AAC GGC ATA AAT	1661
Ile Ala Phe Met Cys Ala Glu Glu Ser Arg Trp Ile Asn Gly Ile Asn	
230 235 240	

ATT CCA GTG GAC GGA GGT TTG GCA TCG ACC TAC GTG TAA GTTCGTGGAC 1710
Ile Pro Val Asp Gly Gly Leu Ala Ser Thr Tyr Val
245 250 255

GCCCTTTGCA CGCGCACTAT ATCTCTATGC AGCAGCTGAA AGCAGCTTTG GTTTTGATCG 1770

GAGGTAGCGG GCGGAAAGGT GCAGAATGTC TAAATAATAA AGGATTCTTG TGAAGCTTTA 1830

GTTGTCCGTA AACGAAAATA AAAATAAAGA GGAATGATAT GAAAGCAAGT AGATCAGTCT 1890

GCACTTTCAA AATAGCTACC CTGGCAGGCG CCATTTATGC AGCGCTGCCA ATGTCAGCTG 1950

CAAACGAT GCAGCTGGAT GTAGGTAGCT CGGATTGGAC GGTGCGTTGG GGACAACACC 2010

~~CTCAAGTATA GCCTTGCCTC TCGCCTGAAT GAGCAAGACT CAAGTCTGAC AAATGCGCCG 2070~~

ACTGTCAATG GTTATATCCG GATATTCAAA GTCAGGGTGA TCGTAACTTT GACCGGGGGC 2130

TTGGTATCCA ATCGTCTCGA TATTCTGGCT GCAG 2164

FIG. 2a:

CTGCAGCCAG GGCTGAAAAG GAGGGATTCA GTGAGGTCAT GAAGGGAGGG GACGGCGCCT	60
GGCTCCAATT GCTCGATGGC GCCGCGATTG AGTGTCTTGG GCGCGGTCTT GGAGAGTTTG	120
GCTAGGGAGA TAAATTTGCT GGCCATGGTG GCGGCCCTTG ATGGGTGGA TGATTTTCTG	180
CATTCTGCAT CATGAAATTC ATGAAATCAT CACTTTTCGG GGGGTGGGTG CACGGGATTG	240
AAGGTTGCTA GGAGAGTGCA TTGCTCGTAA GCCCAGGAAG CACGCGGGTT TCAGGATGGT	300
GCATGGAAT GGCATGAGCT TTGCTGGATA TGATTAGAGA CATTAACAT TTTGGCGGAA	360
TGGAAGCACG ATTCTCGCC CGGTAGAGCG GTAACCGCGA CATTGAGGAC CGTAAAAAGG	420
AAAGAGCATG CAA CTG ACC AAC AAG AAA ATC GTC GTC ACC GGA GTG TCC TCC	472
Met Gln Leu Thr Asn Lys Lys Ile Val Val Thr Gly Val Ser Ser	
1 5 10 15	
GGT ATC GGT GCC GAA ACT GCC CGC GTT CTG CGC TCT CAC GGC GCC ACA	520
Gly Ile Gly Ala Glu Thr Ala Arg Val Leu Arg Ser His Gly Ala Thr	
20 25 30	
GTG ATT GGC GTA GAT CGC AAC ATG CCG AGC CTG ACT CTG GAT GCT TTC	568
Val Ile Gly Val Asp Arg Asn Met Pro Ser Leu Thr Leu Asp Ala Phe	
35 40 45	
GTT CAG GCT GAC CTG AGC CAT CCT GAGGGGAGAG GCGGTTTGCG TATTGGGCGC	622
Val Gln Ala Asp Leu Ser His Pro	
50 55	
ATGCATAAAA ACTGTTGTAA TTCATTAAGC ATTCTGCCGA CATGGAAGCC ATCACAAACG	682
GCATGATGAA CCTGAATCGC CAGCGGCATC AGCACCTTGT CGCCTTGCGT ATAATATTTG	742
CCCATGGACG CACACCGTGG AAACGGATGA AGGCACGAAC CCAGTTGACA TAAGCCTGTT	802
CGGTTTCGTAA ACTGTAATGC AAGTAGCGTA TGCGCTCAGC CAACTGGTCC AGAACCTTGA	862
CCGAACGCAG CGGTGGTAAC GGCGCAGTGG CGGTTTTTCAT GGCTTGTTAT GACTGTTTTT	922
TTGTACAGTC TATGCCTCGG GCATCCAAGC AGCAAGCGCG TTACGCCGTG GGTCGATGTT	982
TGATGTTATG GAGCAGCAAC G ATG TTA CGC AGC AGC AAC GAT GTT ACG CAG	1033
Met Leu Arg Ser Ser Asn Asp Val Thr Gln	
1 5 10	
CAG GGC AGT CGC CCT AAA ACA AAG TTA GGT GGC TCA AGT ATG GGC ATC	1081
Gln Gly Ser Arg Pro Lys Thr Lys Leu Gly Gly Ser Ser Met Gly Ile	
15 20 25	
ATT CGC ACA TGT AGG CTC GGC CCT GAC CAA GTC AAA TCC ATG CGG GCT	1129
Ile Arg Thr Cys Arg Leu Gly Pro Asp Gln Val Lys Ser Met Arg Ala	
30 35 40	
GCT CTT GAT CTT TTC GGT CGT GAG TTC GGA GAC GTA GCC ACC TAC TCC	1177
Ala Leu Asp Leu Phe Gly Arg Glu Phe Gly Asp Val Ala Thr Tyr Ser	
45 50 55	

CAA CAT CAG CCG GAC TCC GAT TAC CTC GGG AAC TTG CTC CGT AGT AAG Gln His Gln Pro Asp Ser Asp Tyr Leu Gly Asn Leu Leu Arg Ser Lys 60 65 70	1225
ACA TTC ATC GCG CTT GCT GCC TTC GAC CAA GAA GCG GTT GTT GGC GCT Thr Phe Ile Ala Leu Ala Ala Phe Asp Gln Glu Ala Val Val Gly Ala 75 80 85 90	1273
CTC GCG GCT TAC GTT CTG CCC AGG TTT GAG CAG CCG CGT AGT GAG ATC Leu Ala Ala Tyr Val Leu Pro Arg Phe Glu Gln Pro Arg Ser Glu Ile 95 100 105	1321
TAT ATC TAT GAT CTC GCA GTC TCC GGC GAG CAC CGG AGG CAG GGC ATT Tyr Ile Tyr Asp Leu Ala Val Ser Gly Glu His Arg Arg Gln Gly Ile 110 115 120	1369
GCC ACC GCG CTC ATC AAT CTC CTC AAG CAT GAG GCC AAC GCG CTT GGT Ala Thr Ala Leu Ile Asn Leu Leu Lys His Glu Ala Asn Ala Leu Gly 125 130 135	1417
GCT TAT GTG ATC TAC GTG CAA GCA GAT TAC GGT GAC GAT CCC GCA GTG Ala Tyr Val Ile Tyr Val Gln Ala Asp Tyr Gly Asp Asp Pro Ala Val 140 145 150	1465
GCT CTC TAT ACA AAG TTG GGC ATA CGG GAA GAA GTG ATG CAC TTT GAT Ala Leu Tyr Thr Lys Leu Gly Ile Arg Glu Glu Val Met His Phe Asp 155 160 165 170	1513
ATC GAC CCA AGT ACC GCC ACC TAA CAATTCGTTT AAGCCGAGAT CGGCTTCCCT Ile Asp Pro Ser Thr Ala Thr 175 177	1567
G ATT GCA TTC ATG TGT GCT GAG GAG TCA CGT TGG ATC AAC GGC ATA AAT Ile Ala Phe Met Cys Ala Glu Glu Ser Arg Trp Ile Asn Gly Ile Asn 228 230 235 240	1616
ATT CCA GTG GAC GGA GGT TTG GCA TCG ACC TAC GTG TAA GTTCGTGGAC Ile Pro Val Asp Gly Gly Leu Ala Ser Thr Tyr Val 245 250 255	1665
GCCCTTTGCA CGCGCACTAT ATCTCTATGC AGCAGCTGAA AGCAGCTTTG GTTTTGATCG	1725
GAGGTAGCGG GCGGAAAGGT GCAGAATGTC TAAATAATAA AGGATTCTTG TGAAGCTTTA	1785
GTTGTCCGTA AACGAAAATA AAAATAAAGA GGAATGATAT GAAAGCAAGT AGATCAGTCT	1845
GCACTTTCAA AATAGCTACC CTGGCAGGCG CCATTTATGC AGCGCTGCCA ATGTCAGCTG	1905
CAAACGAT GCAGCTGGAT GTAGGTAGCT CGGATTGGAC GGTGCGTTGG GGACAACACC	1965
CTCAAGTATA GCCTTGCCCTC TCGCCTGAAT GAGCAAGACT CAAGTCTGAC AAATGCGCCG	2025
ACTGTCAATG GTTATATCCG GATATTCAAA GTCAGGGTGA TCGTAACTTT GACCGGGGGC	2085
TTGGTATCCA ATCGTCTCGA TATTCTGGCT GCAG	2119

FIG. 2b:

CTGCAGCCAG GGCTGAAAAG GAGGGATTCA GTGAGGTCAT GAAGGGAGGG GACGGCGCCT	60
GGCTCCAATT GCTCGATGGC GCCGCGATTG AGTGTCTTGG GCGCGGTCTT GGAGAGTTTCG	120
GCTAGGGAGA TAAATTTGCT GGCCATGGTG GCGGCCCCCTG ATGGGTTGGA TGATTTTCTG	180
CATTCTGCAT CATGAAATTC ATGAAATCAT CACTTTTTCGG GGGGTGGGTG CACGGGATTG	240
AAGGTTGCTA GGAGAGTGCA TTGCTCGTAA GCCCAGGAAG CACGCGGGTT TCAGGATGGT	300
GCATGGAAAT GGCATGAGCT TTGCTGGATA TGATTAGAGA CATTAACTAT TTTGGCGGAA	360
TGGAAGCACG ATTCCTCGCC CGGTAGAGCG GTAACCGCGA CATTGAGGAC CGTAAAAAGG	420
AAAGAGCATG CAA CTG ACC AAC AAG AAA ATC GTC GTC ACC GGA GTG TCC TCC	472
Met Gln Leu Thr Asn Lys Lys Ile Val Val Thr Gly Val Ser Ser	
1 5 10 15	
GGT ATC GGT GCC GAA ACT GCC CGC GTT CTG CGC TCT CAC GGC GCC ACA	520
Gly Ile Gly Ala Glu Thr Ala Arg Val Leu Arg Ser His Gly Ala Thr	
20 25 30	
GTG ATT GGC GTA GAT CGC AAC ATG CCG AGC CTG ACT CTG GAT GCT TTC	568
Val Ile Gly Val Asp Arg Asn Met Pro Ser Leu Thr Leu Asp Ala Phe	
35 40 45	
GTT CAG GCT GAC CTG AGC CAT CCT GAA GGC ATC GATC AAC GGC ATA AAT	617
Val Gln Ala Asp Leu Ser His Pro Glu Gly Ile Asn Gly Ile Asn	
50 55 58 240	
ATT CCA GTG GAC GGA GGT TTG GCA TCG ACC TAC GTG TAA GTTCGTGGAC	666
Ile Pro Val Asp Gly Gly Leu Ala Ser Thr Tyr Val	
245 250 255	
GCCCTTTGCA CGCGCACTAT ATCTCTATGC AGCAGCTGAA AGCAGCTTTG GTTTTGATCG	726
GAGGTAGCGG GCGGAAAGGT GCAGAATGTC TAAATAATAA AGGATTCTTG TGAAGCTTTA	786
GTTGTCCGTA AACGAAAATA AAAATAAAGA GGAATGATAT GAAAGCAAGT AGATCAGTCT	846
GCACTTTCAA AATAGCTACC CTGGCAGGCG CCATTTATGC AGCGCTGCCA ATGTCAGCTG	906
CAAACGAT GCAGCTGGAT GTAGGTAGCT CGGATTGGAC GGTGCGTTGG GGACAACACC	966
CTCAAGTATA GCCTTGCCCTC TCGCCTGAAT GAGCAAGACT CAAGTCTGAC AAATGCGCCG	1026
ACTGTCAATG GTTATATCCG GATATTCAAA GTCAGGGTGA TCGTAACTTT GACCGGGGGC	1086
TTGGTATCCA ATCGTCTCGA TATTCTGGCT GCAG	1120

FIG. 2c:

GAATTCGCG TATCGCCCGG TTCTATCAGC GGGCCGCTTT CGAAAGTCAT GGTGTTAGCC 60
 GGTAGGGTCT TTTTCTTGGC CATGCTTGTT GCCTGAACCT TCGTTGACAT AGGGCAGAGG 120
 TGC GTTTGCGC GCTTCGCTTC GCGATGAACC GCATCGAGAT GCTGAGGTCA GGATTTTTC 180
 TTA ACTCGCG TAAGCATTCT GTCATTTTTT TGGTGGCTTT GAACAGCCTG ATGAAAGGTG 240
 GTCTCGCCCT TTGAGGCCGA TTCTTGGGCG CTTGGCGGCG TCGAAGCGAT GCTCCACTAC 300
 CGATTAAGAT AATTAAAATA AGGAAACCGC ATGGTTTCTT ATGTGAATTT GTCTGGCATA 360
 CTCCAGCTCA AGGGCAATTT TTGGGCTATT GGCTGAGCAG TTGCCTCTAT ATGGTTATTC 420

~~ACAATAACAA TTCACTGCTC ACCACCTCAG CC ATC AGC ATT CTT CCT TTC AAT 473~~

Met Ser Ile Leu Gly Leu Asn
 1 5

GGT GCC CCG GTC GGA GCT GAG CAG CTG GGC TCG GCT CTT GAT CGC ATG 521
 Gly Ala Pro Val Gly Ala Glu Gln Leu Gly Ser Ala Leu Asp Arg Met
 10 15 20

AAG AAG GCG CAC CTG GAG CAG GGG CCT GCA AAC TTG GAG CTG CGT CTG 569
 Lys Lys Ala His Leu Glu Gln Gly Pro Ala Asn Leu Glu Leu Arg Leu
 25 30 35

AGT AGG CTG GAT CGT GCG ATT GCA ATG CTT CTG GAA AAT CGT GAA GCA 617
 Ser Arg Leu Asp Arg Ala Ile Ala Met Leu Leu Glu Asn Arg Glu Ala
 40 45 50 55

ATT GCC GAC GCG GTT TCT GCT GAC TTT GGC AAT CGC AGC CGT GAG CAA 665
 Ile Ala Asp Ala Val Ser Ala Asp Phe Gly Asn Arg Ser Arg Glu Gln
 60 65 70

ACA CTG CTT TGC GAC ATT GCT GGC TCG GTG GCA AGC CTG AAG GAT AGC 713
 Thr Leu Leu Cys Asp Ile Ala Gly Ser Val Ala Ser Leu Lys Asp Ser
 75 80 85

CGC GAG CAC GTG GCC AAA TGG ATG GAG CCC GAA CAT CAC AAG GCG ATG 761
 Arg Glu His Val Ala Lys Trp Met Glu Pro Glu His His Lys Ala Met
 90 95 100

TTT CCA GGG GCG GAG GCA CGC GTT GAG TTT CAG CCG CTG GGT GTC GTT 809
 Phe Pro Gly Ala Glu Ala Arg Val Glu Phe Gln Pro Leu Gly Val Val
 105 110 115

GGG GTC ATT AGT CCC TGG AAC TTC CCT ATC GTA CTG GCC TTT GGG CCG 857
 Gly Val Ile Ser Pro Trp Asn Phe Pro Ile Val Leu Ala Phe Gly Pro
 120 125 130 135

CTG GCC GGC ATA TTC GCA GCA GGT AAT CGC GCC ATG CTC AAG CCG TCC 905
 Leu Ala Gly Ile Phe Ala Ala Gly Asn Arg Ala Met Leu Lys Pro Ser
 140 145 150

GAG CTT ACC CCG CGG ACT TCT GCC CTG CTT GCG GAG CTA ATT GCT CGT 953
 Glu Leu Thr Pro Arg Thr Ser Ala Leu Leu Ala Glu Leu Ile Ala Arg
 155 160 165

TAC TTC GAT GAA ACT GAG CTG ACT ACA GTG CTG GGC GAC GCT GAA GTC	1001
Tyr Phe Asp Glu Thr Glu Leu Thr Thr Val Leu Gly Asp Ala Glu Val	
170 175 180	
GGT GCG CTG TTC AGT GCT CAG CCT TTC GAT CAT CTG ATC TTC ACC GGC	1049
Gly Ala Leu Phe Ser Ala Gln Pro Phe Asp His Leu Ile Phe Thr Gly	
185 190 195	
GGC ACT GCC GTG GCC AAG CAC ATC ATG CGT GCC GCG GCG GAT AAC CTA	1097
Gly Thr Ala Val Ala Lys His Ile Met Arg Ala Ala Ala Asp Asn Leu	
200 205 210 215	
GTG CCC GTT ACC CTG GAA TTG GGT GGC AAA TCG CCG GTG ATC GTT TCC	1145
Val Pro Val Thr Leu Glu Leu Gly Gly Lys Ser Pro Val Ile Val Ser	
220 225 230	
CGC AGT GCA GAT ATG GCG GAC GTT GCA CAA CGG GTG TTG ACG GTG AAA	1193
Arg Ser Ala Asp Met Ala Asp Val Ala Gln Arg Val Leu Thr Val Lys	
235 240 245	
ACC TTC AAT GCC GGG CAA ATC TGT CTG GCA CCG GAC TAT GTG CTG CTG	1241
Thr Phe Asn Ala Gly Gln Ile Cys Leu Ala Pro Asp Tyr Val Leu Leu	
250 255 260	
CCG GAA GGGACAGCAA GCGAACCGGA ATTGCCAGCT GGGGCGCCCT CTGGTAAGGT	1297
Pro Glu	
265	
TGGGAAGCCC TGCAAAGTAA ACTGGATGGC TTTCTTGCCG CCAAGGATCT GATGGCGCAG	1357
GGGATCAAGA TCTGATCAAG AGACAGGATG AGGATCGTTT CGC ATG ATT GAA CAA	1412
Met Ile Glu Gln	
1	
GAT GGA TTG CAC GCA GGT TCT CCG GCC GCT TGG GTG GAG AGG CTA TTC	1460
Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val Glu Arg Leu Phe	
5 10 15 20	
GGC TAT GAC TGG GCA CAA CAG ACA ATC GGC TGC TCT GAT GCC GCC GTG	1508
Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser Asp Ala Ala Val	
25 30 35	
TTC CGG CTG TCA GCG CAG GGG CGC CCG GTT CTT TTT GTC AAG ACC GAC	1556
Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe Val Lys Thr Asp	
40 45 50	
CTG TCC GGT GCC CTG AAT GAA CTG CAG GAC GAG GCA GCG CCG CTA TCG	1604
Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala Ala Arg Leu Ser	
55 60 65	
TGG CTG GCC ACG ACG GGC GTT CCT TGC GCA GCT GTG CTC GAC GTT GTC	1652
Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val Leu Asp Val Val	
70 75 80	

ACT GAA GCG GGA AGG GAC TGG CTG CTA TTG GGC GAA GTG CCG GGG CAG	1700
Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu Val Pro Gly Gln	
85 90 95 100	
GAT CTC CTG TCA TCT CAC CTT GCT CCT GCC GAG AAA GTA TCC ATC ATG	1748
Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys Val Ser Ile Met	
105 110 115	
GCT GAT GCA ATG CGG CGG CTG CAT ACG CTT GAT CCG GCT ACC TGC CCA	1796
Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro Ala Thr Cys Pro	
120 125 130	
TTC GAC CAC CAA GCG AAA CAT CGC ATC GAG CGA GCA CGT ACT CGG ATG	1844
Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala Thr Arg Met	
135 140 145	
GAA GCC GGT CTT GTC GAT CAG GAT GAT CTG GAC GAA GAG CAT CAG GGG	1892
Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu Glu His Gln Gly	
150 155 160	
CTC GCG CCA GCC GAA CTG TTC GCC AGG CTC AAG GCG CGC ATG CCC GAC	1940
Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala Arg Met Pro Asp	
165 170 175 180	
GGC GAG GAT CTC GTC GTG ACC CAT GGC GAT GCC TGC TTG CCG AAT ATC	1988
Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys Leu Pro Asn Ile	
185 190 195	
ATG GTG GAA AAT GGC CGC TTT TCT GGA TTC ATC GAC TGT GGC CGG CTG	2036
Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp Cys Gly Arg Leu	
200 205 210	
GGT GTG GCG GAC CGC TAT CAG GAC ATA GCG TTG GCT ACC CGT GAT ATT	2084
Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala Thr Arg Asp Ile	
215 220 225	
GCT GAA GAG CTT GGC GGC GAA TGG GCT GAC CGC TTC CTC GTG CTT TAC	2132
Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe Leu Val Leu Tyr	
230 235 240	
GGT ATC GCC GCT CCC GAT TCG CAG CGC ATC GCC TTC TAT CGC CTT CTT	2180
Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe Tyr Arg Leu Leu	
245 250 255 260	
GAC GAG TTC TTC TGA GCGGGACTCT GGGGTTCGAA ATGACCGACC AAGCGACGCC	2235
Asp Glu Phe Phe	
264	
CGC CAT GCC AAG CCT GTT CTC GTG CAA AGT CCT GTG GGT GAG TCG AAC	2283
His Ala Lys Pro Val Leu Val Gln Ser Pro Val Gly Glu Ser Asn	
444 445 450 455	
TTG GCG ATG CGC GCA CCC TAC GGA GAA GCG ATC CAC GGA CTG CTC TCT	2331
Leu Ala Met Arg Ala Pro Tyr Gly Glu Ala Ile His Gly Leu Leu Ser	
460 465 470	

GTC CTC CTT TCA ACG GAG TGT TAG AACCGTTGGT AGTGGTTTTG GACGGGCCCC 2385
Val Leu Leu Ser Thr Glu Cys
475 480 481

GGAGCATGCG CTTCTGGGCC CGTTTCTTGA GTATTCATTG GATAGTCACG CGTGGTAGCT 2445

TCGAGCCTGC ACAGCTGATG AGCACCTGG AAGGCGCGCT GTACGCGGAC GACTGGGTTC 2505

ATCTTCGCCA TTCATGACGG AACTCCGTTT CCCAGTACCG CGATGACTAT TTTGCCTCTT 2565

CCGATGTCCG ATTCCACGCC GCCTGACGCT AAGCGGGGGC GGGGGCGCCC GCATCCCAGC 2625

CCAGACAGCA ACAAATGAGT AGGCTCTTGG ATGCCGCGGC GGCTGAGATT GGTAACGGCA 2685

~~ATTTCCTCAA TCTGACGATG GATTCCATTG CCCGTGCTGC CCGCCTCTCA AAAAAACGC 2745~~

TGTACGTCTT GGTGGCGAGC AAGGAAGAAC TCATTTCCTG GTTAGTGGCT CGAGACATGT 2805

CCAACCTTGA GGAATTC 2822

FIG. 2d:

GAATTCGCG TATCGCCCGG TTCTATCAGC GGGCCGCTTT CGAAAGTCAT GGTGTTAGCC 60
 GG TAGGGTCT TTTTCTTGGC CATGCTTGTT GCCTGAACCT TCGTTGACAT AGGGCAGAGG 120
 TGC GTTTGCC GCTTCGCTTC GCGATGAACC GCATCGAGAT GCTGAGGTCA GGATTTTTC 180
 TTA ACTCGCG TAAGCATTCT GTCATTTTTT TGGTGGCTTT GAACAGCCTG ATGAAAGGTG 240
 GTCTCGCCCT TTGAGGCCGA TTCTTGGGCG CTTGGCGGCG TCGAAGCGAT GCTCCACTAC 300
 CGATTAAGAT AATTAATA AGGAAACCGC ATGGTTTCTT ATGTGAATTT GTCTGGCATA 360
 CTCCAGCTCA AGGGCAATTT TTGGGCTATT GGCTGAGCAG TTGCCTCTAT ATGTTATTC 420

~~ACAATAAGAA TTCACTGCTC AGGAGGTCAG CG ATG AGC ATT CTT GGT TTG AAT~~ 473
 Met Ser Ile Leu Gly Leu Asn
 1 5

AGT GCC CCG GTC GGA GCT GAG CAG CTG GGC TCG GCT CTT GAT CGC ATG 521
 Gly Ala Pro Val Gly Ala Glu Gln Leu Gly Ser Ala Leu Asp Arg Met
 10 15 20

AAG AAG GCG CAC CTG GAG CAG GGG CCT GCA AAC TTG GAG CTG CGT CTG 569
 Lys Lys Ala His Leu Glu Gln Gly Pro Ala Asn Leu Glu Leu Arg Leu
 25 30 35

AGT AGG CTG GAT CGT GCG ATT GCA ATG CTT CTG GAA AAT CGT GAA GCA 617
 Ser Arg Leu Asp Arg Ala Ile Ala Met Leu Leu Glu Asn Arg Glu Ala
 40 45 50 55

ATT GCC GAC GCG GTT TCT GCT GAC TTT GGC AAT CGC AGC CGT GAG CAA 665
 Ile Ala Asp Ala Val Ser Ala Asp Phe Gly Asn Arg Ser Arg Glu Gln
 60 65 70

ACA CTG CTT TGC GAC ATT GCT GGC TCG GTG GCA AGC CTG AAG GAT AGC 713
 Thr Leu Leu Cys Asp Ile Ala Gly Ser Val Ala Ser Leu Lys Asp Ser
 75 80 85

CGC GAG CAC GTG GCC AAA TGG ATG GAG CCC GAA CAT CAC AAG GCG ATG 761
 Arg Glu His Val Ala Lys Trp Met Glu Pro Glu His His Lys Ala Met
 90 95 100

TTT CCA GGG GCG GAG GCA CGC GTT GAG TTT CAG CCG CTG GGT GTC GTT 809
 Phe Pro Gly Ala Glu Ala Arg Val Glu Phe Gln Pro Leu Gly Val Val
 105 110 115

GGG GTC ATT AGT CCC TGG AAC TTC CCT ATC GTA CTG GCC TTT GGG CCG 857
 Gly Val Ile Ser Pro Trp Asn Phe Pro Ile Val Leu Ala Phe Gly Pro
 120 125 130 135

CTG GCC GGC ATA TTC GCA GCA GGT AAT CGC GCC ATG CTC AAG CCG TCC 905
 Leu Ala Gly Ile Phe Ala Ala Gly Asn Arg Ala Met Leu Lys Pro Ser
 140 145 150

GAG CTT ACC CCG CGG ACT TCT GCC CTG CTT GCG GAG CTA ATT GCT CGT 953
 Glu Leu Thr Pro Arg Thr Ser Ala Leu Leu Ala Glu Leu Ile Ala Arg
 155 160 165

TAC TTC GAT GAA ACT GAG CTG ACT ACA GTG CTG GGC GAC GCT GAA GTC	1001
Tyr Phe Asp Glu Thr Glu Leu Thr Thr Val Leu Gly Asp Ala Glu Val	
170 175 180	
GGT GCG CTG TTC AGT GCT CAG CCT TTC GAT CAT CTG ATC TTC ACC GGC	1049
Gly Ala Leu Phe Ser Ala Gln Pro Phe Asp His Leu Ile Phe Thr Gly	
185 190 195	
GGC ACT GCC GTG GCC AAG CAC ATC ATG CGT GCC GCG GCG GAT AAC CTA	1097
Gly Thr Ala Val Ala Lys His Ile Met Arg Ala Ala Ala Asp Asn Leu	
200 205 210 215	
GTG CCC GTT ACC CTG GAA TTG GGT GGC AAA TCG CCG GTG ATC GTT TCC	1145
Val Pro Val Thr Leu Glu Leu Gly Gly Lys Ser Pro Val Ile Val Ser	
220 225 230	
CGC AGT GCA GAT ATG GCG GAC GTT GCA CAA CGG GTG TTG ACG GTG AAA	1193
Arg Ser Ala Asp Met Ala Asp Val Ala Gln Arg Val Leu Thr Val Lys	
235 240 245	
ACC TTC AAT GCC GGG CAA ATC TGT CTG GCA CCG GAC TAT GTG CTG GGG	1241
Thr Phe Asn Ala Gly Gln Ile Cys Leu Ala Pro Asp Tyr Val Leu	
250 255 260 262	
GAGAGGCGGT TTGCGTATTG GGCGCATGCA TAAAACTGT TGTAATTCAT TAAGCATTCT	1301
GCCGACATGG AAGCCATCAC AAACGGCATG ATGAACCTGA ATCGCCAGCG GCATCAGCAC	1361
CTTGTCGCCT TGCGTATAAT ATTTGCCCAT GGACGCACAC CGTGGAACG GATGAAGGCA	1421
CGAACCCAGT TGACATAAGC CTGTTCCGTT CGTAACTGT AATGCAAGTA GCGTATGCGC	1481
TCACGCAACT GGTCCAGAAC CTTGACCGAA CGCAGCGGTG GTAACGGCGC AGTGGCGGTT	1541
TTCATGGCTT GTTATGACTG TTTTTTTGTA CAGTCTATGC CTCGGGCATC CAAGCAGCAA	1601
GCGCGTTACG CCGTGGGTCG ATGTTTGATG TTATGGAGCA GCAACG ATG TTA CGC	1656
Met Leu Arg	
1	
AGC AGC AAC GAT GTT ACG CAG CAG GGC AGT CGC CCT AAA ACA AAG TTA	1704
Ser Ser Asn Asp Val Thr Gln Gln Gly Ser Arg Pro Lys Thr Lys Leu	
5 10 15	
GGT GGC TCA AGT ATG GGC ATC ATT CGC ACA TGT AGG CTC GGC CCT GAC	1752
Gly Gly Ser Ser Met Gly Ile Ile Arg Thr Cys Arg Leu Gly Pro Asp	
20 25 30 35	
CAA GTC AAA TCC ATG CGG GCT GCT CTT GAT CTT TTC GGT CGT GAG TTC	1800
Gln Val Lys Ser Met Arg Ala Ala Leu Asp Leu Phe Gly Arg Glu Phe	
40 45 50	
GGA GAC GTA GCC ACC TAC TCC CAA CAT CAG CCG GAC TCC GAT TAC CTC	1848
Gly Asp Val Ala Thr Tyr Ser Gln His Gln Pro Asp Ser Asp Tyr Leu	
55 60 65	
GGG AAC TTG CTC CGT AGT AAG ACA TTC ATC GCG CTT GCT GCC TTC GAC	1896
Gly Asn Leu Leu Arg Ser Lys Thr Phe Ile Ala Leu Ala Ala Phe Asp	
70 75 80	

CAA GAA GCG GTT GTT GGC GCT CTC GCG GCT TAC GTT CTG CCC AGG TTT Gln Glu Ala Val Val Gly Ala Leu Ala Ala Tyr Val Leu Pro Arg Phe 85 90 95	1944
GAG CAG CCG CGT AGT GAG ATC TAT ATC TAT GAT CTC GCA GTC TCC GGC Glu Gln Pro Arg Ser Glu Ile Tyr Ile Tyr Asp Leu Ala Val Ser Gly 100 105 110 115	1992
GAG CAC CGG AGG CAG GGC ATT GCC ACC GCG CTC ATC AAT CTC CTC AAG Glu His Arg Arg Gln Gly Ile Ala Thr Ala Leu Ile Asn Leu Leu Lys 120 125 130	2040
CAT GAG GCC AAC GCG CTT GGT GCT TAT GTG ATC TAC GTG CAA GCA GAT His Glu Ala Asn Ala Leu Gly Ala Tyr Val Ile Tyr Val Gln Ala Asp 135 140 145	2088
TAC GGT GAC GAT CCC GCA GTG GCT CTC TAT ACA AAG TTG GGC ATA CGG Tyr Gly Asp Asp Pro Ala Val Ala Leu Tyr Thr Lys Leu Gly Ile Arg 150 155 160	2136
GAA GAA GTG ATG CAC TTT GAT ATC GAC CCA AGT ACC GCC ACC TAA CAA Glu Glu Val Met His Phe Asp Ile Asp Pro Ser Thr Ala Thr 165 170 175 177	2184
TTCGTTCAAG CCGAGATCGG CTTCCCTG CAA AGT CCT GTG GGT GAG TCG AAC Gln Ser Pro Val Gly Glu Ser Asn 451 455	2236
TTG GCG ATG CGC GCA CCC TAC GGA GAA GCG ATC CAC GGA CTG CTC TCT Leu Ala Met Arg Ala Pro Tyr Gly Glu Ala Ile His Gly Leu Leu Ser 460 465 470	2284
GTC CTC CTT TCA ACG GAG TGT TAG AACCGTTGGT AGTGGTTTTG GACGGGCCCA Val Leu Leu Ser Thr Glu Cys 475 480 481	2338
GGAGCATGCG CTTCTGGGCC CGTTTCTTGA GTATTCATTG GATAGTCACG CGTGGTAGCT	2398
TCGAGCCTGC ACAGCTGATG AGCACCTTGG AAGGCGCGCT GTACGCGGAC GACTGGGTTC	2458
ATCTTCGCCA TTCATGACGG AACTCCGTTC CCCAGTACCG CGATGACTAT TTTGCCTCTT	2518
CCGATGTCCG ATTCCACGCC GCCTGACGCT AAGCGGGGGC GGGGGCGCCC GCATCCCAGC	2578
CCAGACAGCA ACAAATGAGT AGGCTCTTGG ATGCCGCGGC GGCTGAGATT GGTAACGGCA	2638
ATTTTCGTCAA TGTGACGATG GATTCGATTG CCCGTGCTGC CGGCGTCTCA AAAAAACGC	2698
TGTACGTCTT GGTGGCGAGC AAGGAAGAAC TCATTTCCCG GTTAGTGGCT CGAGACATGT	2758
CCAACCTTGA GGAATTC	2775

FIG. 2e:

GAATTCGCG TATCGCCCGG TTCTATCAGC GGGCCGCTTT CGAAAGTCAT GGTGTTAGCC 60
 GGTAGGGTCT TTTTCTTGGC CATGCTTGTT GCCTGAACCT TCGTTGACAT AGGGCAGAGG 120
 TCGGTTTGCC GCTTCGCTTC GCGATGAACC GCATCGAGAT GCTGAGGTCA GGATTTTTC 180
 TTAACTCGCG TAAGCATTCT GTCATTTTTT TGGTGGCTTT GAACAGCCTG ATGAAAGGTG 240
 GTCTCGCCCT TTGAGGCCGA TTCTTGGGCG CTTGGCGGCG TCGAAGCGAT GCTCCACTAC 300
 CGATTAAGAT AATTAAAATA AGGAAACCGC ATGGTTTCTT ATGTGAATTT GTCTGGCATA 360
 CTCCAGCTCA AGGGCAATTT TTGGGCTATT GGCTGAGCAG TTGCCTCTAT ATGGTTATTC 420

~~ACAATAAGAA TTCACCTCTC ACCACCTCAG CG ATG AGC ATT CTT CCT TTC AAT~~ 473
 Met Ser Ile Leu Gly Leu Asn
 1 5

AGT GCC CCG GTC GGA GCT GAG CAG CTG GGC TCG GCT CTT GAT CGC ATG 521
 Gly Ala Pro Val Gly Ala Glu Gln Leu Gly Ser Ala Leu Asp Arg Met
 10 15 20

AAG AAG GCG CAC CTG GAG CAG GGG CCT GCA AAC TTG GAG CTG CGT CTG 569
 Lys Lys Ala His Leu Glu Gln Gly Pro Ala Asn Leu Glu Leu Arg Leu
 25 30 35

AGT AGG CTG GAT CGT GCG ATT GCA ATG CTT CTG GAA AAT CGT GAA GCA 617
 Ser Arg Leu Asp Arg Ala Ile Ala Met Leu Leu Glu Asn Arg Glu Ala
 40 45 50 55

ATT GCC GAC GCG GTT TCT GCT GAC TTT GGC AAT CGC AGC CGT GAG CAA 665
 Ile Ala Asp Ala Val Ser Ala Asp Phe Gly Asn Arg Ser Arg Glu Gln
 60 65 70

ACA CTG CTT TGC GAC ATT GCT GGC TCG GTG GCA AGC CTG AAG GAT AGC 713
 Thr Leu Leu Cys Asp Ile Ala Gly Ser Val Ala Ser Leu Lys Asp Ser
 75 80 85

CGC GAG CAC GTG GCC AAA TGG ATG GAG CCC GAA CAT CAC AAG GCG ATG 761
 Arg Glu His Val Ala Lys Trp Met Glu Pro Glu His His Lys Ala Met
 90 95 100

TTT CCA GGG GCG GAG GCA CGC GTT GAG TTT CAG CCG CTG GGT GTC GTT 809
 Phe Pro Gly Ala Glu Ala Arg Val Glu Phe Gln Pro Leu Gly Val Val
 105 110 115

GGG GTC ATT AGT CCC TGG AAC TTC CCT ATC GTA CTG GCC TTT GGG CCG 857
 Gly Val Ile Ser Pro Trp Asn Phe Pro Ile Val Leu Ala Phe Gly Pro
 120 125 130 135

CTG GCC GGC ATA TTC GCA GCA GGT AAT CGC GCC ATG CTC AAG CCG TCC 905
 Leu Ala Gly Ile Phe Ala Ala Gly Asn Arg Ala Met Leu Lys Pro Ser
 140 145 150

GAG CTT ACC CCG CCG AET TCT GCC CTG CTT GCG GAG CTA ATT GET EGT 953
 Glu Leu Thr Pro Arg Thr Ser Ala Leu Leu Ala Glu Leu Ile Ala Arg
 155 160 165

TAC TTC GAT GAA ACT GAG CTG ACT ACA GTG CTG GGC GAC GCT GAA GTC	1001
Tyr Phe Asp Glu Thr Glu Leu Thr Thr Val Leu Gly Asp Ala Glu Val	
170 175 180	
GGT GCG CTG TTC AGT GCT CAG CCT TTC GAT CAT CTG ATC TTC ACC GGC	1049
Gly Ala Leu Phe Ser Ala Gln Pro Phe Asp His Leu Ile Phe Thr Gly	
185 190 195	
GGC ACT GCC GTG GCC AAG CAC ATC ATG CGT GCC GCG GCG GAT AAC CTA	1097
Gly Thr Ala Val Ala Lys His Ile Met Arg Ala Ala Ala Asp Asn Leu	
200 205 210 215	
GTG CCC GTT ACC CTG GAA TTG GGT GGC AAA TCG CCG GTG ATC GTT TCC	1145
Val Pro Val Thr Leu Glu Leu Gly Gly Lys Ser Pro Val Ile Val Ser	
220 225 230	
CGC AGT GCA GAT ATG GCG GAC GTT GCA CAA CGG GTG TTG ACG GTG AAA	1193
Arg Ser Ala Asp Met Ala Asp Val Ala Gln Arg Val Leu Thr Val Lys	
235 240 245	
ACC TTC AAT GCC GGG CAA ATC TGT CTG GCA CC GTG GGT GAG TCG AAC	1240
Thr Phe Asn Ala Gly Gln Ile Cys Leu Ala Val Gly Glu Ser Asn	
250 255 257 454 455	
TTG GCG ATG CGC GCA CCC TAC GGA GAA GCG ATC CAC GGA CTG CTC TCT	1288
Leu Ala Met Arg Ala Pro Tyr Gly Glu Ala Ile His Gly Leu Leu Ser	
460 465 470	
GTC CTC CTT TCA ACG GAG TGT TAG AACCGTTGGT AGTGGTTTTG GACGGGCCCA	1342
Val Leu Leu Ser Thr Glu Cys	
475 480 481	
GGAGCATGCG CTTCTGGGCC CGTTTCTTGA GTATTCATTG GATAGTCACG CGTGGTAGCT	1402
TCGAGCCTGC ACAGCTGATG AGCACCCTGG AAGGCGCGCT GTACGCGGAC GACTGGGTTC	1462
ATCTTCGCCA TTCATGACGG AACTCCGTTC CCCAGTACCG CGATGACTAT TTTGCCTCTT	1522
CCGATGTCCG ATTCCACGCC GCCTGACGCT AAGCGGGGGC GGGGGCGCCC GCATCCCAGC	1582
CCAGACAGCA ACAAATGAGT AGGCTCTTGG ATGCCGCGGC GGCTGAGATT GGTAACGGCA	1642
ATTCGTCAA TGTGACGATG GATTCGATTG CCCGTGCTGC CGGCGTCTCA AAAAAACGC	1702
TGTACGTCTT GGTGGCGAGC AAGGAAGAAC TCATTTCCCG GTTAGTGGCT CGAGACATGT	1762
CCAACCTTGA GGAATTC	1779

FIG. 2f:

CTGCAGCCGA GCATCGATTG AGCACTTTAC CCAGCTGCGC TGGCTGACCA TTCAGAATGG	60
CCCGCGGCAC TATCCAATCT AAATCGATCT TCGGGCGCCG CGGGCATCAT GCCC GCGGCG	120
CTCGCCTCAT TTCAATCTCT AACTTGATAA AAACAGAGCT GTTCTCCGGT CTTGGTGGAT	180
CAAGGCCAGT CGCGGAGAGT CTCGAAGAGG AGAGTACAGT GAACGCCGAG TCCACATTGC	240
AACCGCAGGC ATCATCATGC TCTGCTCAGC CACGCTACCG CAGTGTGTCG ATTGGTCATC	300
CTCCGGTTGA GGTTACGCAA GACGCTGGAG GTATTGTCCG G ATG CGT TCT CTC GAG	356
Met Arg Ser Leu Glu	
1 5	
<hr/>	
GCG CTT CTT CCC TTC CCG GGT CGA ATT CTT GAG CGT CTC GAG CAT TGG	404
Ala Leu Leu Pro Phe Pro Gly Arg Ile Leu Glu Arg Leu Glu His Trp	
10 15 20	
GCT AAG ACC CGT CCA GAA CAA ACC TGC GTT GCT GCC AGG GCG GCA AAT	452
Ala Lys Thr Arg Pro Glu Gln Thr Cys Val Ala Ala Arg Ala Ala Asn	
25 30 35	
GGG GAA TGG CGT CGT ATC AGC TAC GCG GAA ATG TTC CAC AAC GTC CGC	500
Gly Glu Trp Arg Arg Ile Ser Tyr Ala Glu Met Phe His Asn Val Arg	
40 45 50	
GCC ATC GCA CAG AGC TTG CTT CCT TAC GGA CTA TCG GCA GAG CGT CCG	548
Ala Ile Ala Gln Ser Leu Leu Pro Tyr Gly Leu Ser Ala Glu Arg Pro	
55 60 65	
CTG CTT ATC GTC TCT GGA AAT GAC CTG GAA CAT CTT CAG CTG GCA TTT	596
Leu Leu Ile Val Ser Gly Asn Asp Leu Glu His Leu Gln Leu Ala Phe	
70 75 80 85	
GGG GCT ATG TAT GCG GGC ATT CCC TAT TGC CCG GTG TCT CCT GCT TAT	644
Gly Ala Met Tyr Ala Gly Ile Pro Tyr Cys Pro Val Ser Pro Ala Tyr	
90 95 100	
TCA CTG CTG TCG CAA GAT TTG GCG AAG CTG CGT CAC ATC GTA GGT CTT	692
Ser Leu Leu Ser Gln Asp Leu Ala Lys Leu Arg His Ile Val Gly Leu	
105 110 115	
CTG CAA CCG GGA CTG GTC TTT GCT GCC GAT GCA GCA CCT TTC CAG GGG	740
Leu Gln Pro Gly Leu Val Phe Ala Ala Asp Ala Ala Pro Phe Gln	
120 125 130 132	
ACAGCAAGCG AACCGGAATT GCCAGCTGGG GCGCCCTCTG GTAAGGTTGG GAAGCCCTGC	800
AAAGTAAACT GGATGGCTTT CTGCGGCCA AGGATCTGAT GGCGCAGGGG ATCAAGATCT	860
GATCAAGAGA CAGGATGAGG ATCGTTTCGC ATG ATT GAA CAA GAT GGA TTG CAC	914
Met Ile Glu Gln Asp Gly Leu His	
1 5	
GCA GGT TCT CCG GCC GCT TGG GTG GAG AGG CTA TTC GGC TAT GAC TGG	962
Ala Gly Ser Pro Ala Ala Trp Val Glu Arg Leu Phe Gly Tyr Asp Trp	
10 15 20	

GCA CAA CAG ACA ATC GGC TGC TCT GAT GCC GCC GTG TTC CGG CTG TCA	1010
Ala Gln Gln Thr Ile Gly Cys Ser Asp Ala Ala Val Phe Arg Leu Ser	
25 30 35 40	
GCG CAG GGG CGC CCG GTT CTT TTT GTC AAG ACC GAC CTG TCC GGT GCC	1058
Ala Gln Gly Arg Pro Val Leu Phe Val Lys Thr Asp Leu Ser Gly Ala	
45 50 55	
CTG AAT GAA CTG CAG GAC GAG GCA GCG CGG CTA TCG TGG CTG GCC ACG	1106
Leu Asn Glu Leu Gln Asp Glu Ala Ala Arg Leu Ser Trp Leu Ala Thr	
60 65 70	
ACG GGC GTT CCT TGC GCA GCT GTG CTC GAC GTT GTC ACT GAA GCG GGA	1154
Thr Gly Val Pro Cys Ala Ala Val Leu Asp Val Val Thr Glu Ala Gly	
75 80 85	
AGG GAC TGG CTG CTA TTG GGC GAA GTG CCG GGG CAG GAT CTC CTG TCA	1202
Arg Asp Trp Leu Leu Leu Gly Glu Val Pro Gly Gln Asp Leu Leu Ser	
90 95 100	
TCT CAC CTT GCT CCT GCC GAG AAA GTA TCC ATC ATG GCT GAT GCA ATG	1250
Ser His Leu Ala Pro Ala Glu Lys Val Ser Ile Met Ala Asp Ala Met	
105 110 115 120	
CGG CGG CTG CAT ACG CTT GAT CCG GCT ACC TGC CCA TTC GAC CAC CAA	1298
Arg Arg Leu His Thr Leu Asp Pro Ala Thr Cys Pro Phe Asp His Gln	
125 130 135	
GCG AAA CAT CGC ATC GAG CGA GCA CGT ACT CGG ATG GAA GCC GGT CTT	1346
Ala Lys His Arg Ile Glu Arg Ala Arg Thr Arg Met Glu Ala Gly Leu	
140 145 150	
GTC GAT CAG GAT GAT CTG GAC GAA GAG CAT CAG GGG CTC GCG CCA GCC	1394
Val Asp Gln Asp Asp Leu Asp Glu Glu His Gln Gly Leu Ala Pro Ala	
155 160 165	
GAA CTG TTC GCC AGG CTC AAG GCG CGC ATG CCC GAC GGC GAG GAT CTC	1442
Glu Leu Phe Ala Arg Leu Lys Ala Arg Met Pro Asp Gly Glu Asp Leu	
170 175 180	
GTC GTG ACC CAT GGC GAT GCC TGC TTG CCG AAT ATC ATG GTG GAA AAT	1490
Val Val Thr His Gly Asp Ala Cys Leu Pro Asn Ile Met Val Glu Asn	
185 190 195 200	
GGC CGC TTT TCT GGA TTC ATC GAC TGT GGC CGG CTG GGT GTG GCG GAC	1538
Gly Arg Phe Ser Gly Phe Ile Asp Cys Gly Arg Leu Gly Val Ala Asp	
205 210 215	
CGC TAT CAG GAC ATA GCG TTG GCT ACC CGT GAT ATT GCT GAA GAG CTT	1586
Arg Tyr Gln Asp Ile Ala Leu Ala Thr Arg Asp Ile Ala Glu Glu Leu	
220 225 230	
GGC GGC GAA TGG GCT GAC CGC TTC CTC GTG CTT TAC GGT ATC GCC GCT	1634
Gly Gly Glu Trp Ala Asp Arg Phe Leu Val Leu Tyr Gly Ile Ala Ala	
235 240 245	

CCC GAT TCG CAG CGC ATC GCC TTC TAT CGC CTT CTT GAC GAG TTC TTC 1682
 Pro Asp Ser Gln Arg Ile Ala Phe Tyr Arg Leu Leu Asp Glu Phe Phe
 250 255 260 264

TGAGCGGGAC TCTGGGGTTC GAAATGACCG ACCAAGCGAC GCCCCT GTT TTG CAA 1737
 Val Leu Gln
 563 565

TGG CGG TCG GCG AAA GTT GAT GCG CTG TAT CGT GGT GAA GAT CAA TCC 1785
 Trp Arg Ser Ala Lys Val Asp Ala Leu Tyr Arg Gly Glu Asp Gln Ser
 570 575 580

ATG CTG CGT GAC GAG GCC ACA CTG TGA GTTGGTCAGG GGGGGCTTAC 1832
 Met Leu Arg Asp Glu Ala Thr Leu
 585 589

TCGGCGTTTT CCGACACTGC GTTGGTTGCG GCAGTGCACA CCCCCTGGAT TGATTGCGGG 1892

GGTGCCCTGT CGCTGGTGTC GCCTATCGAC TTAGGGGTAA AGGTCGCTCG CGAAGTTCTG 1952

ATGCGTGCGT CGCTTGAACC ACAAATGGTC GATAGCGTAC TCGCAGGCTC TATGGCTCAA 2012

GCAAGCTTTG ATGCTTACCT GCTCCCGCGG CACATTGGCT TGTACAGCGG TGTTCCCAAG 2072

TCGGTTCCGG CCTTGGGGGT GCAGCGCATT TCGGGCACAG GCTTCGAACT GCTTCGGCAG 2132

GCCGGCGAGC AGATTTCCTCA AGGCGCTGAT CACGTGCTGT GTGTCGCGGG CTGCAG 2188

FIG. 2g:

CTGCAGCCGA GCATCGATTG AGCACTTTAC CCAGCTGCGC TGGCTGACCA TTCAGAATGG 60
 CCCGCGGCAC TATCCAATCT AAATCGATCT TCGGGCGCCG CGGGCATCAT GCCC GCGGCG 120
 CTCGCCTCAT TTCAATCTCT AACTTGATAA AAACAGAGCT GTTCTCCGGT CTTGGTGGAT 180
 CAAGGCCAGT CGCGGAGAGT CTCGAAGAGG AGAGTACAGT GAACGCCGAG TCCACATTGC 240
 AACC GCAGGC ATCATCATGC TCTGCTCAGC CACGCTACCG CAGTGTGTCG ATTGGTCATC 300
 CTCCGTTGA GGTACGCAA GACGCTGGAG GTATTGTCCG G ATG CGT TCT CTC GAG 356
 Met Arg Ser Leu Glu
 1 5

~~GGG CTT CTT CCG TTC CCG GGT CGA ATT CTT GAG CGT CTC GAG CAT TGG 404~~
 Ala Leu Leu Pro Phe Pro Gly Arg Ile Leu Glu Arg Leu Glu His Trp
 10 15 20

GCT AAG ACC CGT CCA GAA CAA ACC TGC GTT GCT GCC AGG GCG GCA AAT 452
 Ala Lys Thr Arg Pro Glu Gln Thr Cys Val Ala Ala Arg Ala Ala Asn
 25 30 35

GGG GAA TGG CGT CGT ATC AGC TAC GCG GAA ATG TTC CAC AAC GTC CGC 500
 Gly Glu Trp Arg Arg Ile Ser Tyr Ala Glu Met Phe His Asn Val Arg
 40 45 50

GCC ATC GCA CAG AGC TTG CTT CCT TAC GGA CTA TCG GCA GAG CGT CCG 548
 Ala Ile Ala Gln Ser Leu Leu Pro Tyr Gly Leu Ser Ala Glu Arg Pro
 55 60 65

CTG CTT ATC GTC TCT GGA AAT GAC CTG GAA CAT CTT CAG CTG GCA TTT 596
 Leu Leu Ile Val Ser Gly Asn Asp Leu Glu His Leu Gln Leu Ala Phe
 70 75 80 85

GGG GCT ATG TAT GCG GGC ATT CCC TAT TGC CCG GTG TCT CCT GCT TAT 644
 Gly Ala Met Tyr Ala Gly Ile Pro Tyr Cys Pro Val Ser Pro Ala Tyr
 90 95 100

TCA CTG CTG TCG CAA GAT TTG GCG AAG CTG CGT CAC ATC GTA GGT CTT 692
 Ser Leu Leu Ser Gln Asp Leu Ala Lys Leu Arg His Ile Val Gly Leu
 105 110 115

CTG CAA CCG GGA CTG GTC TTT GCT GCC GAT GCA GCA CCT TTC CAG GGG 740
 Leu Gln Pro Gly Leu Val Phe Ala Ala Asp Ala Ala Pro Phe Gln
 120 125 130 132

GAGAGGCGGT TTGCGTATTG GGCGCATGCA TAAAACTGT TGTAATTCAT TAAGCATTCT 800

GCCGACATGG AAGCCATCAC AAACGGCATG ATGAACCTGA ATCGCCAGCG GCATCAGCAC 860

CTTGTCGCCT TGC GTATAAT ATTTGCCCAT GGACGCACAC CGTGGAAACG GATGAAGGCA 920

CGAACCCAGT TGACATAAGC CTGTTCCGGTT CGTAACTGT AATGCAAGTA GCGTATGCGC 980

TCACGCAACT GGTCCAGAAC CTGACCGAA CGCAGCGGTG GTAACGGCGC AGTGGCGGTT 1040

TTCATGGCTT GTTATGACTG TTTT TTTGTA CAGTCTATGC CTCGGGCATC CAAGCAGCAA 1100

GCGCGTTACG CCGTGGGTCG ATGTTTGATG TTATGGAGCA GCAACG ATG TTA CGC	1155
Met Leu Arg	
1	
AGC AGC AAC GAT GTT ACG CAG CAG GGC AGT CGC CCT AAA ACA AAG TTA	1203
Ser Ser Asn Asp Val Thr Gln Gln Gly Ser Arg Pro Lys Thr Lys Leu	
5 10 15	
GGT GGC TCA AGT ATG GGC ATC ATT CGC ACA TGT AGG CTC GGC CCT GAC	1251
Gly Gly Ser Ser Met Gly Ile Ile Arg Thr Cys Arg Leu Gly Pro Asp	
20 25 30 35	
CAA GTC AAA TCC ATG CGG GCT GCT CTT GAT CTT TTC GGT CGT GAG TTC	1299
Gln Val Lys Ser Met Arg Ala Ala Leu Asp Leu Phe Gly Arg Glu Phe	
40 45 50	
GGA GAC GTA GCC ACC TAC TCC CAA CAT CAG CCG GAC TCC GAT TAC CTC	1347
Gly Asp Val Ala Thr Tyr Ser Gln His Gln Pro Asp Ser Asp Tyr Leu	
55 60 65	
GGG AAC TTG CTC CGT AGT AAG ACA TTC ATC GCG CTT GCT GCC TTC GAC	1395
Gly Asn Leu Leu Arg Ser Lys Thr Phe Ile Ala Leu Ala Ala Phe Asp	
70 75 80	
CAA GAA GCG GTT GTT GGC GCT CTC GCG GCT TAC GTT CTG CCC AGG TTT	1443
Gln Glu Ala Val Val Gly Ala Leu Ala Ala Tyr Val Leu Pro Arg Phe	
85 90 95	
GAG CAG CCG CGT AGT GAG ATC TAT ATC TAT GAT CTC GCA GTC TCC GGC	1491
Glu Gln Pro Arg Ser Glu Ile Tyr Ile Tyr Asp Leu Ala Val Ser Gly	
100 105 110 115	
GAG CAC CCG AGG CAG GGC ATT GCC ACC GCG CTC ATC AAT CTC CTC AAG	1539
Glu His Arg Arg Gln Gly Ile Ala Thr Ala Leu Ile Asn Leu Leu Lys	
120 125 130	
CAT GAG GCC AAC GCG CTT GGT GCT TAT GTG ATC TAC GTG CAA GCA GAT	1587
His Glu Ala Asn Ala Leu Gly Ala Tyr Val Ile Tyr Val Gln Ala Asp	
135 140 145	
TAC GGT GAC GAT CCC GCA GTG GCT CTC TAT ACA AAG TTG GGC ATA CGG	1635
Tyr Gly Asp Asp Pro Ala Val Ala Leu Tyr Thr Lys Leu Gly Ile Arg	
150 155 160	
GAA GAA GTG ATG CAC TTT GAT ATC GAC CCA AGT ACC GCC ACC TAA CAA	1683
Glu Glu Val Met His Phe Asp Ile Asp Pro Ser Thr Ala Thr	
165 170 175 177	
TTCGTTCAAG CCGAGATCGG CTTCCCCT GTT TTG CAA TGG CGG TCG GCG AAA	1735
Val Leu Gln Trp Arg Ser Ala Lys	
563 565 570	
GTT GAT GCG CTG TAT CGT GGT GAA GAT CAA TCC ATG CTG CGT GAC GAG	1783
Val Asp Ala Leu Tyr Arg Gly Glu Asp Gln Ser Met Leu Arg Asp Glu	
575 580 585	

GCC ACA CTG TGA GTTGGTCAGG GGGGGCTTAC TCGGCGTTTT CCGACACTGC	1835
Ala Thr Leu	
589	
GTGTTGGTTCG GCAGTGCACA CCCCCTGGAT TGATTGCGGG GGTGCCCTGT CGCTGGTGTC	1895
GCCTATCGAC TTAGGGGTAA AGGTCGCTCG CGAAGTTCTG ATGCGTGCGT CGCTTGAACC	1955
ACAAATGGTC GATAGCGTAC TCGCAGGCTC TATGGCTCAA GCAAGCTTTG ATGCTTACCT	2015
GCTCCCGCGG CACATTGGCT TGTACAGCGG TGTTCCCAAG TCGGTTCCGG CCTTGGGGGT	2075
GCAGCGCATT TCGGGCACAG GCTTCGAACT GCTTCGGCAG GCCGGCGAGC AGATTTCCTA	2135
AGGGGCTGAT GAGCTCTCT CTCTCGGGCG CTGGAG	2171

FIG. 2h:

CTGCAGCCGA GCATCGATTG AGCACTTTAC CCAGCTGCGC TGGCTGACCA TTCAGAATGG	60
CCCGCGGCAC TATCCAATCT AAATCGATCT TCGGGCGCCG CGGGCATCAT GCGGCGGGCG	120
CTCGCCTCAT TTCAATCTCT AACTTGATAA AAACAGAGCT GTTCTCCGGT CTTGGTGGAT	180
CAAGGCCAGT CGCGGAGAGT CTCGAAGAGG AGAGTACAGT GAACGCCGAG TCCACATTGC	240
AACCGCAGGC ATCATCATGC TCTGCTCAGC CACGCTACCG CAGTGTGTCG ATTGGTCATC	300
CTCCGGTTGA GGTTACGCAA GACGCTGGAG GTATTGTCCG G ATG CGT TCT CTC GAG	356
Met Arg Ser Leu Glu	
1 5	
CCC CTT CTT CCC TTC CCG GGT CGA ATT CTT GAG CGT CTC GAG CAT TGG	404
Ala Leu Leu Pro Phe Pro Gly Arg Ile Leu Glu Arg Leu Glu His Trp	
10 15 20	
GCT AAG ACC CGT CCA GAA CAA ACC TGC GTT GCT GCC AGG GCG GCA AAT	452
Ala Lys Thr Arg Pro Glu Gln Thr Cys Val Ala Ala Arg Ala Ala Asn	
25 30 35	
GGG GAA TGG CGT CGT ATC AGC TAC GCG GAA ATG TTC CAC AAC GTC CGC	500
Gly Glu Trp Arg Arg Ile Ser Tyr Ala Glu Met Phe His Asn Val Arg	
40 45 50	
GCC ATC GCA CAG AGC TTG CTT CCT TAC GGA CTA TCG GCA GAG CGT CCG	548
Ala Ile Ala Gln Ser Leu Leu Pro Tyr Gly Leu Ser Ala Glu Arg Pro	
55 60 65	
CTG CTT ATC GTC TCT GGA AAT GAC CTG GAA CAT CTT CAG CTG GCA TTT	596
Leu Leu Ile Val Ser Gly Asn Asp Leu Glu His Leu Gln Leu Ala Phe	
70 75 80 85	
GGG GCT ATG TAT GCG GGC ATT CCC TAT TGC CCG GTG TCT CCT GCT TAT	644
Gly Ala Met Tyr Ala Gly Ile Pro Tyr Cys Pro Val Ser Pro Ala Tyr	
90 95 100	
TCA CTG CTG TCG CAA GAT TTG GCG AAG CTG CGT CAC ATC GTA GGT CTT	692
Ser Leu Leu Ser Gln Asp Leu Ala Lys Leu Arg His Ile Val Gly Leu	
105 110 115	
CTG CAA CCG GGA CTG GTC TTT GCT GCC GAT GCA GCA CCT TTC CAG CGC	740
Leu Gln Pro Gly Leu Val Phe Ala Ala Asp Ala Ala Pro Phe Gln Arg	
120 125 130 133	
GCT GTT TTG CAA TGG CGG TCG GCG AAA GTT GAT GCG CTG TAT CGT GGT	788
Ala Val Leu Gln Trp Arg Ser Ala Lys Val Asp Ala Leu Tyr Arg Gly	
562 565 570 575	
GAA GAT CAA TCC ATG CTG CGT GAC GAG GCC ACA CTG TGA GTTGGTCAGG	837
Glu Asp Gln Ser Met Leu Arg Asp Glu Ala Thr Leu	
580 585 589	
GGGGGCTTAC TEGGGGTTTT CCGACACTGC GTTGGTTGCG GCAGTGC GCA CCCCCTGGAT	897
TGATTGCGGG GGTGCCCTGT CGCTGGTGTC GCCTATCGAC TTAGGGGTAA AGGTCGCTCG	957

CGAAGTTCTG ATGCGTGCGT CGCTTGAACC ACAAATGGTC GATAGCGTAC TCGCAGGCTC	1017
TATGGCTCAA GCAAGCTTTG ATGCTTACCT GCTCCCGCGG CACATTGGCT TGTACAGCGG	1077
TGTTCCCAAG TCGGTTCCGG CCTTGGGGGT GCAGCGCATT TCGGGCACAG GCTTCGAACT	1137
GCTTCGGCAG GCCGGCGAGC AGATTTCCTA AGGCGCTGAT CACGTGCTGT GTGTCGCGGG	1197
CTGCAG	1203

FIG. 2i:

GAATTCCCCT GGCACGAAA GGGCGGCAGG CCGCATGGCC ACGGCTGGGC GGTAACATGAT	60
GCTTGCGTTA ATCGTTAACC GTTTGAAATT CCTTGCCAAA TTTCGGCGAG AGAATCATGC	120
GGGTACGCCT TTCCGTGCGC TTTGATCTGC GCTTCCGTGC CTTGAATCAG AAAAATAGTT	180
AATTGACAGA ACTATAGGTT CGCAGTAGCT TTTGCTCACC CACCAAATCC ACAGCACTGG	240
GGTGCACG ATG AAT AGC TAC GAT GGC CGT TGG TCT ACC GTT GAT GTG AAG	290
Met Asn Ser Tyr Asp Gly Arg Trp Ser Thr Val Asp Val Lys	
1 5 10	
GTT GAA GAA GGT ATC GCT TGG GTC ACG CTG AAC CGC CCG GAG AAG CGC	338
Val Glu Glu Gly Ile Ala Trp Val Thr Leu Asn Arg Pro Glu Lys Arg	
15 20 25 30	
AAC GCA ATG AGC CCA ACT CTC AAT CGA GAG ATG GTC GAG GTT CTG GAG	386
Asn Ala Met Ser Pro Thr Leu Asn Arg Glu Met Val Glu Val Leu Glu	
35 40 45	
GTG CTG GAG CAG GAC GCA GAT GCT CGC GTG CTT GTT CTG ACT GGT GCA	434
Val Leu Glu Gln Asp Ala Asp Ala Arg Val Leu Val Leu Thr Gly Ala	
50 55 60	
GGC GAA TCC TGG ACC GCG GGC ATG GAC CTG AAG GAG TAT TTC CGC GAG	482
Gly Glu Ser Trp Thr Ala Gly Met Asp Leu Lys Glu Tyr Phe Arg Glu	
65 70 75	
ACC GAT GCT GGC CCC GAA ATT CTG CAA GAG AAG ATT CGT CGGGGACAGC	531
Thr Asp Ala Gly Pro Glu Ile Leu Gln Glu Lys Ile Arg	
80 85 90 91	
AAGCGAACCG GAATTGCCAG CTGGGGCGCC CTCTGGTAAG GTTGGGAAGC CCTGCAAAGT	591
AAACTGGATG GCTTTCTTGC CGCCAAGGAT CTGATGGCGC AGGGGATCAA GATCTGATCA	651
AGAGACAGGA TGAGGATCGT TTCGC ATG ATT GAA CAA GAT GGA TTG CAC GCA	703
Met Ile Glu Gln Asp Gly Leu His Ala	
1 5	
GGT TCT CCG GCC GCT TGG GTG GAG AGG CTA TTC GGC TAT GAC TGG GCA	751
Gly Ser Pro Ala Ala Trp Val Glu Arg Leu Phe Gly Tyr Asp Trp Ala	
10 15 20 25	
CAA CAG ACA ATC GGC TGC TCT GAT GCC GCC GTG TTC CGG CTG TCA GCG	799
Gln Gln Thr Ile Gly Cys Ser Asp Ala Ala Val Phe Arg Leu Ser Ala	
30 35 40	
CAG GGG CGC CCG GTT CTT TTT GTC AAG ACC GAC CTG TCC GGT GCC CTG	847
Gln Gly Arg Pro Val Leu Phe Val Lys Thr Asp Leu Ser Gly Ala Leu	
45 50 55	
AAT GAA CTG CAG GAC GAG GCA GCG CGG CTA TCG TGG CTG GCC ACG ACG	895
Asn Glu Leu Gln Asp Glu Ala Ala Arg Leu Ser Trp Leu Ala Thr Thr	
60 65 70	

GGC	GTT	CCT	TGC	GCA	GCT	GTG	CTC	GAC	GTT	GTC	ACT	GAA	GCG	GGA	AGG	943
Gly	Val	Pro	Cys	Ala	Ala	Val	Leu	Asp	Val	Val	Thr	Glu	Ala	Gly	Arg	
75						80					85					
GAC	TGG	CTG	CTA	TTG	GGC	GAA	GTG	CCG	GGG	CAG	GAT	CTC	CTG	TCA	TCT	991
Asp	Trp	Leu	Leu	Leu	Gly	Glu	Val	Pro	Gly	Gln	Asp	Leu	Leu	Ser	Ser	
90					95				100						105	
CAC	CTT	GCT	CCT	GCC	GAG	AAA	GTA	TCC	ATC	ATG	GCT	GAT	GCA	ATG	CGG	1039
His	Leu	Ala	Pro	Ala	Glu	Lys	Val	Ser	Ile	Met	Ala	Asp	Ala	Met	Arg	
				110					115					120		
CGG	CTG	CAT	ACG	CTT	GAT	CCG	GCT	ACC	TGC	CCA	TTC	GAC	CAC	CAA	GCG	1087
Arg	Leu	His		Leu	Asp	Pro	Ala	Thr	Cys	Pro	Phe	Asp	His	Gln	Ala	
			125					130					135			
AAA	CAT	CGC	ATC	GAG	CGA	GCA	CGT	ACT	CGG	ATG	GAA	GCC	GGT	CTT	GTC	1135
Lys	His	Arg	Ile	Glu	Arg	Ala	Arg	Thr	Arg	Met	Glu	Ala	Gly	Leu	Val	
		140					145					150				
GAT	CAG	GAT	GAT	CTG	GAC	GAA	GAG	CAT	CAG	GGG	CTC	GCG	CCA	GCC	GAA	1183
Asp	Gln	Asp	Asp	Leu	Asp	Glu	Glu	His	Gln	Gly	Leu	Ala	Pro	Ala	Glu	
155						160					165					
CTG	TTC	GCC	AGG	CTC	AAG	GCG	CGC	ATG	CCC	GAC	GGC	GAG	GAT	CTC	GTC	1231
Leu	Phe	Ala	Arg	Leu	Lys	Ala	Arg	Met	Pro	Asp	Gly	Glu	Asp	Leu	Val	
170					175					180					185	
GTG	ACC	CAT	GGC	GAT	GCC	TGC	TTG	CCG	AAT	ATC	ATG	GTG	GAA	AAT	GGC	1279
Val	Thr	His	Gly	Asp	Ala	Cys	Leu	Pro	Asn	Ile	Met	Val	Glu	Asn	Gly	
				190					195					200		
CGC	TTT	TCT	GGA	TTC	ATC	GAC	TGT	GGC	CGG	CTG	GGT	GTG	GCG	GAC	CGC	1327
Arg	Phe	Ser	Gly	Phe	Ile	Asp	Cys	Gly	Arg	Leu	Gly	Val	Ala	Asp	Arg	
			205					210					215			
TAT	CAG	GAC	ATA	GCG	TTG	GCT	ACC	CGT	GAT	ATT	GCT	GAA	GAG	CTT	GGC	1375
Tyr	Gln	Asp	Ile	Ala	Leu	Ala	Thr	Arg	Asp	Ile	Ala	Glu	Glu	Leu	Gly	
		220					225					230				
GGC	GAA	TGG	GCT	GAC	CGC	TTC	CTC	GTG	CTT	TAC	GGT	ATC	GCC	GCT	CCC	1423
Gly	Glu	Trp	Ala	Asp	Arg	Phe	Leu	Val	Leu	Tyr	Gly	Ile	Ala	Ala	Pro	
235						240					245					
GAT	TCG	CAG	CGC	ATC	GCC	TTC	TAT	CGC	CTT	CTT	GAC	GAG	TTC	TTC	TGA	1471
Asp	Ser	Gln	Arg	Ile	Ala	Phe	Tyr	Arg	Leu	Leu	Asp	Glu	Phe	Phe		
250					255				260				264			
GC	GGG	ACTCT	GGGGTTCGAA	ATGACCGACC	AAGCGACGCC	CC	GAG	CAG	GGC	ATG						1525
										Glu	Gln	Gly	Met			
										255						
AAG	CAG	TTC	CTT	GAC	GAG	AAA	AGC	ATC	AAG	CCG	GGC	TTG	CAG	ACC	TAC	1573
Lys	Gln	Phe	Leu	Asp	Glu	Lys	Ser	Ile	Lys	Pro	Gly	Leu	Gln	Thr	Tyr	
260						265					270					

AAG CGC TGA TAAATGCGCC GGGGCCCTCG CTGCGCCCCC GGCCTTCCAA TAATGACAAT 1632
Lys Arg
275 276

AATGAGGAGT GCCCAATGTT TCACGTGCCC CTGCTTATTG GTGGTAAGCC TTGTTTACGCA 1692

TCTGATGAGC GCACCTTCGA GCGTCGTAGC CCGCTGACCG GAGAAGTGGT ATCGCGCGTC 1752

GCTGCTGCCA GTTTGGAAGA TGCGGACGCC GCAGTGGCCG CTGCACAGGC TCGGTTTCCT 1812

GAATGGGCGG CGCTTGCTCC GAGCGAACGC CGTGCCCGAC TGCTGCGAGC GCGGATCTT 1872

CTAGAGGACC GTTCTTCCGA GTTCACCGCC GCAGCGAGTG AACTGGCGC AGCGGGAAAC 1932

~~TGGTATGGGT TTAACGTTA CCTGGCGGCG GGCATGTTGC GGCGAATTC 1981~~

FIG. 2j:

GAATTCCTT	GGCGACGAAA	GGGCGGCAGG	CCGCATGGCC	ACGGCTGGGC	GGTAACTGAT	60
GCTTGCGTTA	ATCGTTAACC	GTTTGAAATT	CCTTGCCAAA	TTTCGGCGAG	AGAATCATGC	120
GGGTACGCCT	TTCCGTGCGC	TTTGATCTGC	GCTTCCGTGC	CTTGAATCAG	AAAAATAGTT	180
AATTGACAGA	ACTATAGGTT	CGCAGTAGCT	TTTGCTCACC	CACCAAATCC	ACAGCACTGG	240
GGTGACAG	ATG AAT AGC TAC GAT GGC CGT TGG TCT ACC GTT GAT GTG AAG	290				
	Met Asn Ser Tyr Asp Gly Arg Trp Ser Thr Val Asp Val Lys					
	1 5 10					
GTT GAA GAA GGT ATC GCT TGG GTC ACG CTG AAC CGC CCG GAG AAG CGC	338					
Val Glu Glu Gly Ile Ala Trp Val Thr Leu Asn Arg Pro Glu Lys Arg						
15 20 25 30						
AAC GCA ATG AGC CCA ACT CTC AAT CGA GAG ATG GTC GAG GTT CTG GAG	386					
Asn Ala Met Ser Pro Thr Leu Asn Arg Glu Met Val Glu Val Leu Glu						
35 40 45						
GTG CTG GAG CAG GAC GCA GAT GCT CGC GTG CTT GTT CTG ACT GGT GCA	434					
Val Leu Glu Gln Asp Ala Asp Ala Arg Val Leu Val Leu Thr Gly Ala						
50 55 60						
GGC GAA TCC TGG ACC GCG GGC ATG GAC CTG AAG GAG TAT TTC CGC GAG	482					
Gly Glu Ser Trp Thr Ala Gly Met Asp Leu Lys Glu Tyr Phe Arg Glu						
65 70 75						
ACC GAT GCT GGC CCC GAA ATT CTG CAA GAG AAG ATT CGT CGGGGGAGAG	531					
Thr Asp Ala Gly Pro Glu Ile Leu Gln Glu Lys Ile Arg						
80 85 90 91						
GCGGTTTGCG TATTGGGCGC ATGCATAAAA ACTGTTGTAA TTCATTAAGC ATTCTGCCGA	591					
CATGGAAGCC ATCACAAACG GCATGATGAA CCTGAATCGC CAGCGGCATC AGCACCTTGT	651					
CGCCTTGCGT ATAATATTTG CCCATGGACG CACACCGTGG AAACGGATGA AGGCACGAAC	711					
CCAGTTGACA TAAGCCTGTT CGGTTTCGTAA ACTGTAATGC AAGTAGCGTA TGCGCTCACG	771					
CAACTGGTCC AGAACCTTGA CCGAACGCAG CGGTGGTAAC GGCGCAGTGG CGGTTTTTCAT	831					
GGCTTGTTAT GACTGTTTTT TTGTACAGTC TATGCCTCGG GCATCCAAGC AGCAAGCGCG	891					
TTACGCCGTG GGTGATGTT TGATGTTATG GAGCAGCAAC G ATG TTA CGC AGC AGC	947					
Met Leu Arg Ser Ser						
1 5						
AAC GAT GTT ACG CAG CAG GGC AGT CGC CCT AAA ACA AAG TTA GGT GGC	995					
Asn Asp Val Thr Gln Gln Gly Ser Arg Pro Lys Thr Lys Leu Gly Gly						
10 15 20						
TCA AGT ATG GGC ATC ATT CGC ACA TGT AGG CTC GGC CCT GAC CAA GTC	1043					
Ser Ser Met Gly Ile Ile Arg Thr Cys Arg Leu Gly Pro Asp Gln Val						
25 30 35						

AAA TCC ATG CGG GCT GCT CTT GAT CTT TTC GGT CGT GAG TTC GGA GAC	1091
Lys Ser Met Arg Ala Ala Leu Asp Leu Phe Gly Arg Glu Phe Gly Asp	
40 45 50	
GTA GCC ACC TAC TCC CAA CAT CAG CCG GAC TCC GAT TAC CTC GGG AAC	1139
Val Ala Thr Tyr Ser Gln His Gln Pro Asp Ser Asp Tyr Leu Gly Asn	
55 60 65	
TTG CTC CGT AGT AAG ACA TTC ATC GCG CTT GCT GCC TTC GAC CAA GAA	1187
Leu Leu Arg Ser Lys Thr Phe Ile Ala Leu Ala Ala Phe Asp Gln Glu	
70 75 80 85	
GCG GTT GTT GGC GCT CTC GCG GCT TAC GTT CTG CCC AGG TTT GAG CAG	1235
Ala Val Val Gly Ala Leu Ala Ala Tyr Val Leu Pro Arg Phe Glu Gln	
90 95 100	
CCG CGT AGT GAG ATC TAT ATC TAT GAT CTC GCA GTC TCC GGC GAG CAC	1283
Pro Arg Ser Glu Ile Tyr Ile Tyr Asp Leu Ala Val Ser Gly Glu His	
105 110 115	
CGG AGG CAG GGC ATT GCC ACC GCG CTC ATC AAT CTC CTC AAG CAT GAG	1331
Arg Arg Gln Gly Ile Ala Thr Ala Leu Ile Asn Leu Leu Lys His Glu	
120 125 130	
GCC AAC GCG CTT GGT GCT TAT GTG ATC TAC GTG CAA GCA GAT TAC GGT	1379
Ala Asn Ala Leu Gly Ala Tyr Val Ile Tyr Val Gln Ala Asp Tyr Gly	
135 140 145	
GAC GAT CCC GCA GTG GCT CTC TAT ACA AAG TTG GGC ATA CGG GAA GAA	1427
Asp Asp Pro Ala Val Ala Leu Tyr Thr Lys Leu Gly Ile Arg Glu Glu	
150 155 160 165	
GTG ATG CAC TTT GAT ATC GAC CCA AGT ACC GCC ACC TAA CAATTCGTTC	1476
Val Met His Phe Asp Ile Asp Pro Ser Thr Ala Thr	
170 175 177	
AAGCCGAGAT CGGCTTCCCC GAG CAG GGC ATG AAG CAG TTC CTT GAC GAG	1526
Glu Gln Gly Met Lys Gln Phe Leu Asp Glu	
255 260	
AAA AGC ATC AAG CCG GGC TTG CAG ACC TAC AAG CGC TGA TAAATGCGCC	1575
Lys Ser Ile Lys Pro Gly Leu Gln Thr Tyr Lys Arg	
265 270 275 276	
GGGGCCCTCG CTGCGCCCC GGCCTTCCAA TAATGACAAT AATGAGGAGT GCCCAATGTT	1635
TCACGTGCCC CTGCTTATTG GTGGTAAGCC TTGTTTCTAGCA TCTGATGAGC GCACCTTCGA	1695
GCGTCGTAGC CCGCTGACCG GAGAAGTGGT ATCGCGCGTC GCTGCTGCCA GTTTGGAAGA	1755
TGCGGACGCC GCAGTGGCCG CTGCACAGGC TGCGTTTCTT GAATGGGCGG CGCTTGCTCC	1815
GAGCGAACGC CGTGCCCGAC TGCTGCGAGC GCGGATCTT CTAGAGGACC GTTCTTCCGA	1875

GTTCACCGCC GCAGCGAGTG AACTGGCGC AGCGGGAAAC TGGTATGGGT TTAACGTTA 1935

CCTGGCGGCG GGCATGTTGC GGGGAATTC 1964

FIG. 2k:

GAATTC	CCCT	GGCGACGAAA	GGGCGGCAGG	CCGCATGGCC	ACGGCTGGGC	GGTAACTGAT	60
GCTTGC	GTTA	ATCGTTAACC	GTTTGAAATT	CCTTGCCAAA	TTTCGGCGAG	AGAATCATGC	120
GGGTAC	GCCT	TTCCGTGCGC	TTTGATCTGC	GCTTCCGTGC	CTTGAATCAG	AAAAATAGTT	180
AATTGACAGA	ACTATAGGTT	CGCAGTAGCT	TTTGCTCACC	CACCAAATCC	ACAGCACTGG		240
GGTGCACG	ATG AAT AGC TAC GAT GGC CGT TGG TCT ACC GTT GAT GTG AAG						290
	Met Asn Ser Tyr Asp Gly Arg Trp Ser Thr Val Asp Val Lys						
	1 5 10						
GTT GAA GAA GGT ATC GCT TGG GTC ACG CTG AAC CGC CCG GAG AAG CGC							338
Val Glu Glu Gly Ile Ala Trp Val Thr Leu Asn Arg Pro Glu Lys Arg							
15 20 25 30							
AAC GCA ATG AGC CCA ACT CTC AAT CGA GAG ATG GTC GAG GTT CTG GAG							386
Asn Ala Met Ser Pro Thr Leu Asn Arg Glu Met Val Glu Val Leu Glu							
	35 40 45						
GTG CTG GAG CAG GAC GCA GAT GCT CGC GTG CTT GTT CTG ACT GGT GCA							434
Val Leu Glu Gln Asp Ala Asp Ala Arg Val Leu Val Leu Thr Gly Ala							
	50 55 60						
GGC GAA TCC TGG ACC GCG GGC ATG GAC CTG AAG GAG TAT TTC CGC GAG							482
Gly Glu Ser Trp Thr Ala Gly Met Asp Leu Lys Glu Tyr Phe Arg Glu							
	65 70 75						
ACC GAT GCT GGC CCC GAA ATT CTG CAA GAG AAG ATT CGT CGC GAG CAG							530
Thr Asp Ala Gly Pro Glu Ile Leu Gln Glu Lys Ile Arg Arg Glu Gln							
	80 85 90 92 255						
GGC ATG AAG CAG TTC CTT GAC GAG AAA AGC ATC AAG CCG GGC TTG CAG							578
Gly Met Lys Gln Phe Leu Asp Glu Lys Ser Ile Lys Pro Gly Leu Gln							
	260 265 270						
ACC TAC AAG CGC TGA TAAATGCGCC GGGGCCCTCG CTGCGCCCCC GGCCTTCCAA							633
Thr Tyr Lys Arg							
	275 276						
TAATGACAAT AATGAGGAGT GCCCAATGTT TCACGTGCCC CTGCTTATTG GTGGTAAGCC							693
TTGTTCAGCA TCTGATGAGC GCACCTTCGA GCGTCGTAGC CCGCTGACCG GAGAAGTGGT							753
ATCGCGCGTC GCTGCTGCCA GTTTGGAAGA TGCGGACGCC GCAGTGGCCG CTGCACAGGC							813
TGCGTTTCCT GAATGGGCGG CGCTTGCTCC GAGCGAACGC CGTGCCCGAC TGCTGCGAGC							873
GGCGGATCTT CTAGAGGACC GTTCTTCCGA GTTCACCGCC GCAGCGAGTG AACTGGCGC							933
AGCGGGAAAC TGGTATGGGT TTAACGTTTA CCTGGCGGCG GGCATGTTGC GGGGAATTC							992

FIG. 21:

GAATTCCAAT AATGACAATA ATGAGGAGTG CCCA ATG TTT CAC GTG CCC CTG CTT	55
Met Phe His Val Pro Leu Leu	
1 5	
ATT GGT GGT AAG CCT TGT TCA GCA TCT GAT GAG CGC ACC TTC GAG CGT	103
Ile Gly Gly Lys Pro Cys Ser Ala Ser Asp Glu Arg Thr Phe Glu Arg	
10 15 20	
CGT AGC CCG CTG ACC GGA GAA GTG GTA TCG CGC GTC GCT GCT GCC AGT	151
Arg Ser Pro Leu Thr Gly Glu Val Val Ser Arg Val Ala Ala Ala Ser	
25 30 35	
TTG GAA GAT GCG GAC GCC GCA GTG GCC GCT GCA CAG GCT GCG TTT CCT	199
Leu Glu Asp Ala Asp Ala Ala Val Ala Ala Ala Gln Ala Ala Phe Pro	
40 45 50 55	
GAA TGG GCG GCG CTT GCT CCG AGC GAA CGC CGT GCC CGA CTG CTG CGA	247
Glu Trp Ala Ala Leu Ala Pro Ser Glu Arg Arg Ala Arg Leu Leu Arg	
60 65 70	
GCG GCG GAT CTT CTA GAG GAC CGT TCT TCC GAG TTC ACC GCC GCA GCG	295
Ala Ala Asp Leu Leu Glu Asp Arg Ser Ser Glu Phe Thr Ala Ala Ala	
75 80 85	
AGT GAA ACT GGC GCA GCG GGA AAC TGG TAT GGG TTT AAC GTT TAC CTG	343
Ser Glu Thr Gly Ala Ala Gly Asn Trp Tyr Gly Phe Asn Val Tyr Leu	
90 95 100	
GCG GCG GGC ATG TTG CGG GAA GCC GCG GCC ATG ACC ACA CAG ATT CAG	391
Ala Ala Gly Met Leu Arg Glu Ala Ala Ala Met Thr Thr Gln Ile Gln	
105 110 115	
GGC GAT GTC ATT CCG TCC AAT GTG CCC GGT AGC TTT GCC ATG GCG GTT	439
Gly Asp Val Ile Pro Ser Asn Val Pro Gly Ser Phe Ala Met Ala Val	
120 125 130 135	
CGA CAG CCA TGT GGC GTG GTG CTC GGT ATT GCG CCT TGG AAT GCT CCG	487
Arg Gln Pro Cys Gly Val Val Leu Gly Ile Ala Pro Trp Asn Ala Pro	
140 145 150	
GTA ATC CTT GGC GTA CGG GCT GTT GCG ATG CCG TTG GCA TGC GGC AAT	535
Val Ile Leu Gly Val Arg Ala Val Ala Met Pro Leu Ala Cys Gly Asn	
155 160 165	
ACC GTG GTG TTG AAA AGC TCT GAG CTG AGT CCC TTT ACC CAT CGC CTG	583
Thr Val Val Leu Lys Ser Ser Glu Leu Ser Pro Phe Thr His Arg Leu	
170 175 180	
ATT GGT CAG GTG TTG CAT GAT GCT GGT CTG GGG GAT GGC GTG GTG AAT	631
Ile Gly Gln Val Leu His Asp Ala Gly Leu Gly Asp Gly Val Val Asn	
185 190 195	
GTC ATC AGC AAT GCC CCG CAA GAC GCT CCT GCG GTG GTG GAG CGA CTG	679
Val Ile Ser Asn Ala Pro Gln Asp Ala Pro Ala Val Val Glu Arg Leu	
200 205 210 215	

ATT GCA AAT CCT GCG GTA CGT CGA GTG AAC TTC ACC GGT TCG ACC CAC	727
Ile Ala Asn Pro Ala Val Arg Arg Val Asn Phe Thr Gly Ser Thr His	
220 225 230	
GTT GGA CGG ATC ATT GGT GAG CTG TCT GCG CGT CAT CTG AAG CCT GCT	775
Val Gly Arg Ile Ile Gly Glu Leu Ser Ala Arg His Leu Lys Pro Ala	
235 240 245	
GTG CTG GAA TTA GGT GGT AAG GCT CCG TTC TTG GTC TTG GAC GAT GCC	823
Val Leu Glu Leu Gly Gly Lys Ala Pro Phe Leu Val Leu Asp Asp Ala	
250 255 260	
GAC CTC GAT GCG GCG GTC GAA GCG GCG GCC TTT GGT GCC TAC TTC AAT	871
Asp Leu Asp Ala Ala Val Glu Ala Ala Ala Phe Gly Ala Tyr Phe Asn	
265 270 275	
CAG GGT CAA ATC TGC ATG TCC ACT GAG CGT CTG ATT GTG ACA GCA GTC	919
Gln Gly Gln Ile Cys Met Ser Thr Glu Arg Leu Ile Val Thr Ala Val	
280 285 290 295	
GCA GAC GCC TTT GTT GAA AAG CTG GCG AGG AAG GTC GCC ACA CTG CGT	967
Ala Asp Ala Phe Val Glu Lys Leu Ala Arg Lys Val Ala Thr Leu Arg	
300 305 310	
GCT GGC GAT CCT AAT GAT CCG CAA TCG GTC TTG GGT TCG TTG ATT GAT	1015
Ala Gly Asp Pro Asn Asp Pro Gln Ser Val Leu Gly Ser Leu Ile Asp	
315 320 325	
GCC AAT GCA GGT CAA CGC ATC CAG GTT CTG GTC GAT GAT GCG CTC GGG	1063
Ala Asn Ala Gly Gln Arg Ile Gln Val Leu Val Asp Asp Ala Leu	
330 335 340 342	
GACAGCAAGC GAACCGGAAT TGCCAGCTGG GCGGCCCTCT GGTAAGGTTG GGAAGCCCTG	1123
CAAAGTAAAC TGGATGGCTT TCTTGCCGCC AAGGATCTGA TGGCGCAGGG GATCAAGATC	1183
TGATCAAGAG ACAGGATGAG GATCGTTTCG C ATG ATT GAA CAA GAT GGA TTG	1235
Met Ile Glu Gln Asp Gly Leu	
1 5	
CAC GCA GGT TCT CCG GCC GCT TGG GTG GAG AGG CTA TTC GGC TAT GAC	1283
His Ala Gly Ser Pro Ala Ala Trp Val Glu Arg Leu Phe Gly Tyr Asp	
10 15 20	
TGG GCA CAA CAG ACA ATC GGC TGC TCT GAT GCC GCC GTG TTC CGG CTG	1331
Trp Ala Gln Gln Thr Ile Gly Cys Ser Asp Ala Ala Val Phe Arg Leu	
25 30 35	
TCA GCG CAG GGG CGC CCG GTT CTT TTT GTC AAG ACC GAC CTG TCC GGT	1379
Ser Ala Gln Gly Arg Pro Val Leu Phe Val Lys Thr Asp Leu Ser Gly	
40 45 50 55	
GCC CTG AAT GAA CTG CAG GAC GAG GCA GCG CGG CTA TCG TGG CTG GCC	1427
Ala Leu Asn Glu Leu Gln Asp Glu Ala Ala Arg Leu Ser Trp Leu Ala	
60 65 70	

ACG	ACG	GGC	GTT	CCT	TGC	GCA	GCT	GTG	CTC	GAC	GTT	GTC	ACT	GAA	GCG	1475
Thr	Thr	Gly	Val	Pro	Cys	Ala	Ala	Val	Leu	Asp	Val	Val	Thr	Glu	Ala	
		75						80					85			
GGA	AGG	GAC	TGG	CTG	CTA	TTG	GGC	GAA	GTG	CCG	GGG	CAG	GAT	CTC	CTG	1523
Gly	Arg	Asp	Trp	Leu	Leu	Leu	Gly	Glu	Val	Pro	Gly	Gln	Asp	Leu	Leu	
		90					95					100				
TCA	TCT	CAC	CTT	GCT	CCT	GCC	GAG	AAA	GTA	TCC	ATC	ATG	GCT	GAT	GCA	1571
Ser	Ser	His	Leu	Ala	Pro	Ala	Glu	Lys	Val	Ser	Ile	Met	Ala	Asp	Ala	
	105					110					115					
ATG	CGG	CGG	CTG	CAT	ACG	CTT	GAT	CCG	GCT	ACC	TGC	CCA	TTC	GAC	CAC	1619
Met	Arg	Arg	Leu	His	Thr	Leu	Asp	Pro	Ala	Thr	Cys	Pro	Phe	Asp	His	
120						125				130					135	
CAA	GCG	AAA	CAT	CGC	ATC	GAG	CGA	GCA	CGT	ACT	CGG	ATG	GAA	GCC	GGT	1667
Gln	Ala	Lys	His	Arg	Ile	Glu	Arg	Ala	Arg	Thr	Arg	Met	Glu	Ala	Gly	
			140						145					150		
CTT	GTC	GAT	CAG	GAT	GAT	CTG	GAC	GAA	GAG	CAT	CAG	GGG	CTC	GCG	CCA	1715
Leu	Val	Asp	Gln	Asp	Asp	Leu	Asp	Glu	Glu	His	Gln	Gly	Leu	Ala	Pro	
		155						160					165			
GCC	GAA	CTG	TTC	GCC	AGG	CTC	AAG	GCG	CGC	ATG	CCC	GAC	GGC	GAG	GAT	1763
Ala	Glu	Leu	Phe	Ala	Arg	Leu	Lys	Ala	Arg	Met	Pro	Asp	Gly	Glu	Asp	
	170						175					180				
CTC	GTC	GTG	ACC	CAT	GGC	GAT	GCC	TGC	TTG	CCG	AAT	ATC	ATG	GTG	GAA	1811
Leu	Val	Val	Thr	His	Gly	Asp	Ala	Cys	Leu	Pro	Asn	Ile	Met	Val	Glu	
	185					190					195					
AAT	GGC	CGC	TTT	TCT	GGA	TTC	ATC	GAC	TGT	GGC	CGG	CTG	GGT	GTG	GCG	1859
Asn	Gly	Arg	Phe	Ser	Gly	Phe	Ile	Asp	Cys	Gly	Arg	Leu	Gly	Val	Ala	
200					205					210					215	
GAC	CGC	TAT	CAG	GAC	ATA	GCG	TTG	GCT	ACC	CGT	GAT	ATT	GCT	GAA	GAG	1907
Asp	Arg	Tyr	Gln	Asp	Ile	Ala	Leu	Ala	Thr	Arg	Asp	Ile	Ala	Glu	Glu	
			220						225					230		
CTT	GGC	GGC	GAA	TGG	GCT	GAC	CGC	TTC	CTC	GTG	CTT	TAC	GGT	ATC	GCC	1955
Leu	Gly	Gly	Glu	Trp	Ala	Asp	Arg	Phe	Leu	Val	Leu	Tyr	Gly	Ile	Ala	
		235						240					245			
GCT	CCC	GAT	TCG	CAG	CGC	ATC	GCC	TTC	TAT	CGC	CTT	CTT	GAC	GAG	TTC	2003
Ala	Pro	Asp	Ser	Gln	Arg	Ile	Ala	Phe	Tyr	Arg	Leu	Leu	Asp	Glu	Phe	
	250					255						260				
TTC	TGA	GCGGGACTCT	GGGGTTCGAA	ATGACCGACC	AAGCGACGCC	CG	GCC	CAG								2057
Phe								Ala	Gln							
264								421								
CGC	GTC	GAT	TCG	GGC	ATT	TGC	CAT	ATC	AAT	GGA	CCG	ACT	GTG	CAT	GAC	2105
Arg	Val	Asp	Ser	Gly	Ile	Cys	His	Ile	Asn	Gly	Pro	Thr	Val	His	Asp	
	425					430						435				

GAG GCT CAG ATG CCA TTC GGT GGG GTG AAG TCC AGC GGC TAC GGC AGC 2153
Glu Ala Gln Met Pro Phe Gly Gly Val Lys Ser Ser Gly Tyr Gly Ser
440 445 450

TTC GGC AGT CGA GCA TCG ATT GAG CAC TTT ACC CAG CTG CGC TGG CTG 2201
Phe Gly Ser Arg Ala Ser Ile Glu His Phe Thr Gln Leu Arg Trp Leu
455 460 465 470

ACC ATT CAG AAT GGC CCG CGG CAC TAT CCA ATC TAA ATCGATCTTC 2247
Thr Ile Gln Asn Gly Pro Arg His Tyr Pro Ile
475 480 481

GGGCGCCGCG GGCATCATGC CCGCGGCGCT CGCCTCATTT CAATCTCTAA CTTGATAAAA 2307

~~ACACAGCTGT TCTCCGGTCT TGGTGGATCA AGGCCAGTCG CGGAGAGTCT CGAAGAGGAG 2367~~

AGTACAGTGA ACGCCGAGTC CACATTGCAA CCGCAGGCAT CATCATGCTC TGCTCAGCCA 2427

CGCTACCGCA GTGTGTCGAT TGGTCATCCT CCGGTTGAGG TTACGCAAGA CGCTGGAGGT 2487

ATTGTCCGGA TGC GTTCTCT CGAGGCGCTT CTTCCCTTCC CGGGTGAAT TC 2539

FIG. 2m:

GAATTCCAAT AATGACAATA ATGAGGAGTG CCCA ATG TTT CAC GTG CCC CTG CTT	55
Met Phe His Val Pro Leu Leu	
1 5	
ATT GGT GGT AAG CCT TGT TCA GCA TCT GAT GAG CGC ACC TTC GAG CGT	103
Ile Gly Gly Lys Pro Cys Ser Ala Ser Asp Glu Arg Thr Phe Glu Arg	
10 15 20	
CGT AGC CCG CTG ACC GGA GAA GTG GTA TCG CGC GTC GCT GCT GCC AGT	151
Arg Ser Pro Leu Thr Gly Glu Val Val Ser Arg Val Ala Ala Ala Ser	
25 30 35	
TTG GAA GAT GCG GAC GCC GCA GTG GCC GCT GCA CAG GCT GCG TTT CCT	199
Leu Glu Asp Ala Asp Ala Ala Val Ala Ala Ala Gln Ala Ala Phe Pro	
40 45 50 55	
GAA TGG GCG GCG CTT GCT CCG AGC GAA CGC CGT GCC CGA CTG CTG CGA	247
Glu Trp Ala Ala Leu Ala Pro Ser Glu Arg Arg Ala Arg Leu Leu Arg	
60 65 70	
GCG GCG GAT CTT CTA GAG GAC CGT TCT TCC GAG TTC ACC GCC GCA GCG	295
Ala Ala Asp Leu Leu Glu Asp Arg Ser Ser Glu Phe Thr Ala Ala Ala	
75 80 85	
AGT GAA ACT GGC GCA GCG GGA AAC TGG TAT GGG TTT AAC GTT TAC CTG	343
Ser Glu Thr Gly Ala Ala Gly Asn Trp Tyr Gly Phe Asn Val Tyr Leu	
90 95 100	
GCG GCG GGC ATG TTG CGG GAA GCC GCG GCC ATG ACC ACA CAG ATT CAG	391
Ala Ala Gly Met Leu Arg Glu Ala Ala Ala Met Thr Thr Gln Ile Gln	
105 110 115	
GGC GAT GTC ATT CCG TCC AAT GTG CCC GGT AGC TTT GCC ATG GCG GTT	439
Gly Asp Val Ile Pro Ser Asn Val Pro Gly Ser Phe Ala Met Ala Val	
120 125 130 135	
CGA CAG CCA TGT GGC GTG GTG CTC GGT ATT GCG CCT TGG AAT GCT CCG	487
Arg Gln Pro Cys Gly Val Val Leu Gly Ile Ala Pro Trp Asn Ala Pro	
140 145 150	
GTA ATC CTT GGC GTA CGG GCT GTT GCG ATG CCG TTG GCA TGC GGC AAT	535
Val Ile Leu Gly Val Arg Ala Val Ala Met Pro Leu Ala Cys Gly Asn	
155 160 165	
ACC GTG GTG TTG AAA AGC TCT GAG CTG AGT CCC TTT ACC CAT CGC CTG	583
Thr Val Val Leu Lys Ser Ser Glu Leu Ser Pro Phe Thr His Arg Leu	
170 175 180	
ATT GGT CAG GTG TTG CAT GAT GCT GGT CTG GGG GAT GGC GTG GTG AAT	631
Ile Gly Gln Val Leu His Asp Ala Gly Leu Gly Asp Gly Val Val Asn	
185 190 195	
GTC ATC AGC AAT GCC CCG CAA GAC GCT CCT GCG GTG GTG GAG CGA CTG	679
Val Ile Ser Asn Ala Pro Gln Asp Ala Pro Ala Val Val Glu Arg Leu	
200 205 210 215	

ATT GCA AAT CCT GCG GTA CGT CGA GTG AAC TTC ACC GGT TCG ACC CAC	727
Ile Ala Asn Pro Ala Val Arg Arg Val Asn Phe Thr Gly Ser Thr His	
220 225 230	
GTT GGA CGG ATC ATT GGT GAG CTG TCT GCG CGT CAT CTG AAG CCT GCT	775
Val Gly Arg Ile Ile Gly Glu Leu Ser Ala Arg His Leu Lys Pro Ala	
235 240 245	
GTG CTG GAA TTA GGT GGT AAG GCT CCG TTC TTG GTC TTG GAC GAT GCC	823
Val Leu Glu Leu Gly Gly Lys Ala Pro Phe Leu Val Leu Asp Asp Ala	
250 255 260	
GAC CTC GAT GCG GCG GTC GAA GCG GCG GCC TTT GGT GCC TAC TTC AAT	871
Asp Leu Asp Ala Ala Val Glu Ala Ala Phe Gly Ala Tyr Phe Asn	
265 270 275	
CAG GGT CAA ATC TGC ATG TCC ACT GAG CGT CTG ATT GTG ACA GCA GTC	919
Gln Gly Gln Ile Cys Met Ser Thr Glu Arg Leu Ile Val Thr Ala Val	
280 285 290 295	
GCA GAC GCC TTT GTT GAA AAG CTG GCG AGG AAG GTC GCC ACA CTG CGT	967
Ala Asp Ala Phe Val Glu Lys Leu Ala Arg Lys Val Ala Thr Leu Arg	
300 305 310	
GCT GGC GAT CCT AAT GAT CCG CAA TCG GTC TTG GGT TCG TTG ATT GAT	1015
Ala Gly Asp Pro Asn Asp Pro Gln Ser Val Leu Gly Ser Leu Ile Asp	
315 320 325	
GCC AAT GCA GGT CAA CGC ATC CAG GTGGGGAGAG GCGGTTTGCG TATTGGGCGC	1069
Ala Asn Ala Gly Gln Arg Ile Gln	
330 335	
ATGCATAAAA ACTGTTGTAA TTCATTAAGC ATTCTGCCGA CATGGAAGCC ATCACAAACG	1129
GCATGATGAA CCTGAATCGC CAGCGGCATC AGCACCTTGT CGCCTTGCGT ATAATATTTG	1189
CCCATGGACG CACACCGTGG AAACGGATGA AGGCACGAAC CCAGTTGACA TAAGCCTGTT	1249
CGGTTCGTAA ACTGTAATGC AAGTAGCGTA TGCGCTCACG CAACTGGTCC AGAACCTTGA	1309
CCGAACGCAG CGGTGGTAAC GGCAGCTGG CGGTTTTTCAT GGCTTGTTAT GACTGTTTTT	1369
TTGTACAGTC TATGCCTCGG GCATCCAAGC AGCAAGCGCG TTACGCCGTG GGTCGATGTT	1429
TGATGTTATG GAGCAGCAAC G ATG TTA CGC AGC AGC AAC GAT GTT ACG CAG	1480
Met Leu Arg Ser Ser Asn Asp Val Thr Gln	
1 5 10	
CAG GGC AGT CGC CCT AAA ACA AAG TTA GGT GGC TCA AGT ATG GGC ATC	1528
Gln Gly Ser Arg Pro Lys Thr Lys Leu Gly Gly Ser Ser Met Gly Ile	
15 20 25	
ATT CGC ACA TGT AGG CTC GGC CCT GAC CAA GTC AAA TCC ATG CGG GCT	1576
Ile Arg Thr Cys Arg Leu Gly Pro Asp Gln Val Lys Ser Met Arg Ala	
30 35 40	

GCT CTT GAT CTT TTC GGT CGT GAG TTC GGA GAC GTA GCC ACC TAC TCC	1624
Ala Leu Asp Leu Phe Gly Arg Glu Phe Gly Asp Val Ala Thr Tyr Ser	
45 50 55	
CAA CAT CAG CCG GAC TCC GAT TAC CTC GGG AAC TTG CTC CGT AGT AAG	1672
Gln His Gln Pro Asp Ser Asp Tyr Leu Gly Asn Leu Leu Arg Ser Lys	
60 65 70	
ACA TTC ATC GCG CTT GCT GCC TTC GAC CAA GAA GCG GTT GTT GGC GCT	1720
Thr Phe Ile Ala Leu Ala Ala Phe Asp Gln Glu Ala Val Val Gly Ala	
75 80 85 90	
CTC GCG GCT TAC GTT CTG CCC AGG TTT GAG CAG CCG CGT AGT GAG ATC	1768
Leu Ala Ala Tyr Val Leu Pro Arg Phe Glu Gln Pro Arg Ser Glu Ile	
95 100 105	
TAT ATC TAT GAT CTC GCA GTC TCC GGC GAG CAC CGG AGG CAG GGC ATT	1816
Tyr Ile Tyr Asp Leu Ala Val Ser Gly Glu His Arg Arg Gln Gly Ile	
110 115 120	
GCC ACC GCG CTC ATC AAT CTC CTC AAG CAT GAG GCC AAC GCG CTT GGT	1864
Ala Thr Ala Leu Ile Asn Leu Leu Lys His Glu Ala Asn Ala Leu Gly	
125 130 135	
GCT TAT GTG ATC TAC GTG CAA GCA GAT TAC GGT GAC GAT CCC GCA GTG	1912
Ala Tyr Val Ile Tyr Val Gln Ala Asp Tyr Gly Asp Asp Pro Ala Val	
140 145 150	
GCT CTC TAT ACA AAG TTG GGC ATA CGG GAA GAA GTG ATG CAC TTT GAT	1960
Ala Leu Tyr Thr Lys Leu Gly Ile Arg Glu Glu Val Met His Phe Asp	
155 160 165 170	
ATC GAC CCA AGT ACC GCC ACC TAA CAATTCGTTC AAGCCGAGAT CGGCTTCCCA	2014
Ile Asp Pro Ser Thr Ala Thr	
175 177	
A TTG GCC CAG CGC GTC GAT TCG GGC ATT TGC CAT ATC AAT GGA CCG ACT	2063
Leu Ala Gln Arg Val Asp Ser Gly Ile Cys His Ile Asn Gly Pro Thr	
420 425 430 435	
GTG CAT GAC GAG GCT CAG ATG CCA TTC GGT GGG GTG AAG TCC AGC GGC	2111
Val His Asp Glu Ala Gln Met Pro Phe Gly Gly Val Lys Ser Ser Gly	
440 445 450	
TAC GGC AGC TTC GGC AGT CGA GCA TCG ATT GAG CAC TTT ACC CAG CTG	2159
Tyr Gly Ser Phe Gly Ser Arg Ala Ser Ile Glu His Phe Thr Gln Leu	
455 460 465	
CGC TGG CTG ACC ATT CAG AAT GGC CCG CGG CAC TAT CCA ATC TAA	2204
Arg Trp Leu Thr Ile Gln Asn Gly Pro Arg His Tyr Pro Ile	
470 475 480 481	
ATCGATCTTC GGGCGCCGCG GGCATCATGC CCGCGGCGCT CGCCTCATTT CAATCTCTAA	2264
CTTGATAAAA ACAGAGCTGT TCTCCGGTCT TGGTGATCA AGGCCAGTCG CGGAGAGTCT	2324

CGAAGAGGAG AGTACAGTGA ACGCCGAGTC CACATTGCAA CCGCAGGCAT CATCATGCTC	2384
TGCTCAGCCA CGCTACCGCA GTGTGTCGAT TGGTCATCCT CCGGTTGAGG TTACGCAAGA	2444
CGCTGGAGGT ATTGTCCGGA TGC GTTCTCT CGAGGCGCTT CTTCCCTTCC CGGGTGAAT	2504
TC	2506

FIG. 2n:

GAATTCCAAT AATGACAATA ATGAGGAGTG CCCA ATG TTT CAC GTG CCC CTG CTT	55
Met Phe His Val Pro Leu Leu	
1 5	
ATT GGT GGT AAG CCT TGT TCA GCA TCT GAT GAG CGC ACC TTC GAG CGT	103
Ile Gly Gly Lys Pro Cys Ser Ala Ser Asp Glu Arg Thr Phe Glu Arg	
10 15 20	
CGT AGC CCG CTG ACC GGA GAA GTG GTA TCG CGC GTC GCT GCT GCC AGT	151
Arg Ser Pro Leu Thr Gly Glu Val Val Ser Arg Val Ala Ala Ala Ser	
25 30 35	
TTG GAA GAT GCG GAC GCC GCA GTG GCC GCT GCA CAG GCT GCG TTT CCT	199
Leu Glu Asp Ala Asp Ala Val Ala Ala Gln Ala Ala Phe Pro	
40 45 50 55	
GAA TGG GCG GCG CTT GCT CCG AGC GAA CGC CGT GCC CGA CTG CTG CGA	247
Glu Trp Ala Ala Leu Ala Pro Ser Glu Arg Arg Ala Arg Leu Leu Arg	
60 65 70	
GCG GCG GAT CTT CTA GAG GAC CGT TCT TCC GAG TTC ACC GCC GCA GCG	295
Ala Ala Asp Leu Leu Glu Asp Arg Ser Ser Glu Phe Thr Ala Ala Ala	
75 80 85	
AGT GAA ACT GGC GCA GCG GGA AAC TGG TAT GGG TTT AAC GTT TAC CTG	343
Ser Glu Thr Gly Ala Ala Gly Asn Trp Tyr Gly Phe Asn Val Tyr Leu	
90 95 100	
GCG GCG GGC ATG TTG CGG GAA GCC GCG GCC ATG ACC ACA CAG ATT CAG	391
Ala Ala Gly Met Leu Arg Glu Ala Ala Ala Met Thr Thr Gln Ile Gln	
105 110 115	
GGC GAT GTC ATT CCG TCC AAT GTG CCC GGT AGC TTT GCC ATG GCG GTT	439
Gly Asp Val Ile Pro Ser Asn Val Pro Gly Ser Phe Ala Met Ala Val	
120 125 130 135	
CGA CAG CCA TGT GGC GTG GTG CTC GGT ATT GCG CCT TGG AAT GCT CCG	487
Arg Gln Pro Cys Gly Val Val Leu Gly Ile Ala Pro Trp Asn Ala Pro	
140 145 150	
GTA ATC CTT GGC GTA CGG GCT GTT GCG ATG CCG TTG GCA TGC GGC AAT	535
Val Ile Leu Gly Val Arg Ala Val Ala Met Pro Leu Ala Cys Gly Asn	
155 160 165	
ACC GTG GTG TTG AAA AGC TCT GAG CTG AGT CCC TTT ACC CAT CGC CTG	583
Thr Val Val Leu Lys Ser Ser Glu Leu Ser Pro Phe Thr His Arg Leu	
170 175 180	
ATT GGT CAG GTG TTG CAT GAT GCT GGT CTG GGG GAT GGC GTG GTG AAT	631
Ile Gly Gln Val Leu His Asp Ala Gly Leu Gly Asp Gly Val Val Asn	
185 190 195	
GTC ATC AGC AAT GCC CCG CAA GAC GCT CCT GCG GTG GTG GAG CGA CTG	679
Val Ile Ser Asn Ala Pro Gln Asp Ala Pro Ala Val Val Glu Arg Leu	
200 205 210 215	

ATT GCA AAT CCT GCG GTA CGT CGA GTG AAC TTC ACC GGT TCG ACC CAC	727
Ile Ala Asn Pro Ala Val Arg Arg Val Asn Phe Thr Gly Ser Thr His	
220 225 230	
GTT GGA CGG ATC ATT GGT GAG CTG TCT GCG CGT CAT CTG AAG CCT GCT	775
Val Gly Arg Ile Ile Gly Glu Leu Ser Ala Arg His Leu Lys Pro Ala	
235 240 245	
GTG CTG GAA TTA GGT GGT AAG GCT CCG TTC TTG GTC TTG GAC GAT GCC	823
Val Leu Glu Leu Gly Gly Lys Ala Pro Phe Leu Val Leu Asp Asp Ala	
250 255 260	
GAC CTC GAT GCG GCG GTC GAA GCG GCG GCC TTT GGT GCC TAC TTC AAT	871
Asp Leu Asp Ala Ala Val Glu Ala Ala Ala Phe Gly Ala Tyr Phe Asn	
265 270 275	
CAG GGT CAA ATC TGC ATG TCC ACT GAG CGT CTG ATT GTG ACA GCA GTC	919
Gln Gly Gln Ile Cys Met Ser Thr Glu Arg Leu Ile Val Thr Ala Val	
280 285 290 295	
GCA GAC GCC TTT GTT GAA AAG CTG GCG AGG AAG GTC GCC ACA CTG CGT	967
Ala Asp Ala Phe Val Glu Lys Leu Ala Arg Lys Val Ala Thr Leu Arg	
300 305 310	
GCT GGC GAT CCT AAT GAT CCG CAA TCG GTC TTG GGT TCG TTG ATT GAT	1015
Ala Gly Asp Pro Asn Asp Pro Gln Ser Val Leu Gly Ser Leu Ile Asp	
315 320 325	
GCC AAT GCA GGT CAA CGC ATC CAG GTT CTG GTC GAT GAT GCG CTC GCA	1063
Ala Asn Ala Gly Gln Arg Ile Gln Val Leu Val Asp Asp Ala Leu Ala	
330 335 340	
AAA GGC GCG CAATGGAA TTG GCC CAG CGC GTC GAT TCG GGC ATT TGC CAT	1113
Lys Gly Ala Leu Ala Gln Arg Val Asp Ser Gly Ile Cys His	
345 346 420 425 430	
ATC AAT GGA CCG ACT GTG CAT GAC GAG GCT CAG ATG CCA TTC GGT GGG	1161
Ile Asn Gly Pro Thr Val His Asp Glu Ala Gln Met Pro Phe Gly Gly	
435 440 445	
GTG AAG TCC AGC GGC TAC GGC AGC TTC GGC AGT CGA GCA TCG ATT GAG	1209
Val Lys Ser Ser Gly Tyr Gly Ser Phe Gly Ser Arg Ala Ser Ile Glu	
450 455 460	
CAC TTT ACC CAG CTG CGC TGG CTG ACC ATT CAG AAT GGC CCG CGG CAC	1257
His Phe Thr Gln Leu Arg Trp Leu Thr Ile Gln Asn Gly Pro Arg His	
465 470 475	
TAT CCA ATC TAA ATCGATCTTC GGGCGCCGCG GGCATCATGC CCGCGGCGCT	1309
Tyr Pro Ile	
480 481	
CGCCTCATTT CAATCTCTAA CTTGATAAAA ACAGAGCTGT TCTCCGGTCT TGGTGGATCA	1369
AGGCCAGTCG CGGAGAGTCT CGAAGAGGAG AGTACAGTGA ACGCCGAGTC CACATTGCAA	1429

CCGCAGGCAT CATCATGCTC TGCTCAGCCA CGCTACCGCA GTGTGTCGAT TGGTCATCCT 1489

CCGGTTGAGG TTACGCAAGA CGCTGGAGGT ATTGTCCGGA TCGGTTCTCT CGAGGCGCTT 1549

CTTCCCTTCC CGGGTGGAAT TC 1571

FIG. 2o:

GAATTCCGCG GTCGGCGAAA GTTGATGCGC TGTATCGTGG TGAAGATCAA TCCATGCTGC	60
GTGACGAGGC CACACT GTG AGT TGG TCA GGG GGG GCT TAC TCG GCG TTT TCC	112
Met Ser Trp Ser Gly Gly Ala Tyr Ser Ala Phe Ser	
1 5 10	
GAC ACT GCG TTG GTT GCG GCA GTG CGC ACC CCC TGG ATT GAT TGC GGG	160
Asp Thr Ala Leu Val Ala Ala Val Arg Thr Pro Trp Ile Asp Cys Gly	
15 20 25	
GGT GCC CTG TCG CTG GTG TCG CCT ATC GAC TTA GGG GTA AAG GTC GCT	208
Gly Ala Leu Ser Leu Val Ser Pro Ile Asp Leu Gly Val Lys Val Ala	
30 35 40	
CGC GAA GTT CTG ATG CCT GCG TCG CTT GAA CCA CAA ATG GTC GAT AGC	256
Arg Glu Val Leu Met Arg Ala Ser Leu Glu Pro Gln Met Val Asp Ser	
45 50 55 60	
GTA CTC GCA GGC TCT ATG GCT CAA GCA AGC TTT GAT GCT TAC CTG CTC	304
Val Leu Ala Gly Ser Met Ala Gln Ala Ser Phe Asp Ala Tyr Leu Leu	
65 70 75	
CCG CGG CAC ATT GGC TTG TAC AGC GGT GTT CCC AAG TCG GTT CCG GCC	352
Pro Arg His Ile Gly Leu Tyr Ser Gly Val Pro Lys Ser Val Pro Ala	
80 85 90	
TTG GGG GTG CAG CGC ATT TGC GGC ACA GGC TTC GAA CTG CTT CGG CAG	400
Leu Gly Val Gln Arg Ile Cys Gly Thr Gly Phe Glu Leu Leu Arg Gln	
95 100 105	
GCC GGC GAG CAG ATT TCC CAA GGC GCT GAT CAC GTG CTG TGT GTC GCG	448
Ala Gly Glu Gln Ile Ser Gln Gly Ala Asp His Val Leu Cys Val Ala	
110 115 120	
GCA GAG TCC ATG TCG CGT AAC CCC ATC GCG TCG TAT ACA CAC CGG GGC	496
Ala Glu Ser Met Ser Arg Asn Pro Ile Ala Ser Tyr Thr His Arg Gly	
125 130 135 140	
GGG TTC CGC CTC GGT GCG CCC GTT GAG TTC AAG GAT TTT TTG TGG GAG	544
Gly Phe Arg Leu Gly Ala Pro Val Glu Phe Lys Asp Phe Leu Trp Glu	
145 150 155	
GCA TTG TTT GAT CCT GCT CCA GGA CTC GAC ATG ATC GCT ACC GCA GAA	592
Ala Leu Phe Asp Pro Ala Pro Gly Leu Asp Met Ile Ala Thr Ala Glu	
160 165 170	
AAC CTG GGGACAGCAA GCGAACCGGA ATTGCCAGCT GGGGCGCCCT CTGGTAAGGT	648
Asn Leu	
174	
TGGGAAGCCC TGCAAAGTAA ACTGGATGGC TTTCTTGCCG CCAAGGATCT GATGGCGCAG	708
GGGATCAAGA TCTGATCAAG AGACAGGATG AGGATCGTTT CGC ATG ATT GAA CAA	763
Met Ile Glu Gln	
1	

GAT	GGA	TTG	CAC	GCA	GGT	TCT	CCG	GCC	GCT	TGG	GTG	GAG	AGG	CTA	TTC	811
Asp	Gly	Leu	His	Ala	Gly	Ser	Pro	Ala	Ala	Trp	Val	Glu	Arg	Leu	Phe	
5					10					15					20	
GGC	TAT	GAC	TGG	GCA	CAA	CAG	ACA	ATC	GGC	TGC	TCT	GAT	GCC	GCC	GTG	859
Gly	Tyr	Asp	Trp	Ala	Gln	Gln	Thr	Ile	Gly	Cys	Ser	Asp	Ala	Ala	Val	
				25					30						35	
TTC	CGG	CTG	TCA	GCG	CAG	GGG	CGC	CCG	GTT	CTT	TTT	GTC	AAG	ACC	GAC	907
Phe	Arg	Leu	Ser	Ala	Gln	Gly	Arg	Pro	Val	Leu	Phe	Val	Lys	Thr	Asp	
			40					45					50			
CTG	TCC	GGT	GCC	CTG	AAT	GAA	CTG	CAG	GAC	GAG	GCA	GCG	CGG	CTA	TCG	955
Leu	Ser	Gly	Ala	Leu	Asn	Glu	Leu	Gln	Asp	Glu	Ala	Ala	Arg	Leu	Ser	
		55					60					65				
TGG	CTG	GCC	ACG	ACG	GGC	GTT	CCT	TGC	GCA	GCT	GTG	CTC	GAC	GTT	GTC	1003
Trp	Leu	Ala	Thr	Thr	Gly	Val	Pro	Cys	Ala	Ala	Val	Leu	Asp	Val	Val	
	70					75					80					
ACT	GAA	GCG	GGA	AGG	GAC	TGG	CTG	CTA	TTG	GGC	GAA	GTG	CCG	GGG	CAG	1051
Thr	Glu	Ala	Gly	Arg	Asp	Trp	Leu	Leu	Leu	Gly	Glu	Val	Pro	Gly	Gln	
85					90					95					100	
GAT	CTC	CTG	TCA	TCT	CAC	CTT	GCT	CCT	GCC	GAG	AAA	GTA	TCC	ATC	ATG	1099
Asp	Leu	Leu	Ser	Ser	His	Leu	Ala	Pro	Ala	Glu	Lys	Val	Ser	Ile	Met	
				105					110					115		
GCT	GAT	GCA	ATG	CGG	CGG	CTG	CAT	ACG	CTT	GAT	CCG	GCT	ACC	TGC	CCA	1147
Ala	Asp	Ala	Met	Arg	Arg	Leu	His	Thr	Leu	Asp	Pro	Ala	Thr	Cys	Pro	
			120					125					130			
TTC	GAC	CAC	CAA	GCG	AAA	CAT	CGC	ATC	GAG	CGA	GCA	CGT	ACT	CGG	ATG	1195
Phe	Asp	His	Gln	Ala	Lys	His	Arg	Ile	Glu	Arg	Ala	Arg	Thr	Arg	Met	
		135					140					145				
GAA	GCC	GGT	CTT	GTC	GAT	CAG	GAT	GAT	CTG	GAC	GAA	GAG	CAT	CAG	GGG	1243
Glu	Ala	Gly	Leu	Val	Asp	Gln	Asp	Asp	Leu	Asp	Glu	Glu	His	Gln	Gly	
	150					155					160					
CTC	GCG	CCA	GCC	GAA	CTG	TTC	GCC	AGG	CTC	AAG	GCG	CGC	ATG	CCC	GAC	1291
Leu	Ala	Pro	Ala	Glu	Leu	Phe	Ala	Arg	Leu	Lys	Ala	Arg	Met	Pro	Asp	
165					170					175					180	
GGC	GAG	GAT	CTC	GTC	GTG	ACC	CAT	GGC	GAT	GCC	TGC	TTG	CCG	AAT	ATC	1339
Gly	Glu	Asp	Leu	Val	Val	Thr	His	Gly	Asp	Ala	Cys	Leu	Pro	Asn	Ile	
				185					190						195	
ATG	GTG	GAA	AAT	GGC	CGC	TTT	TCT	GGA	TTC	ATC	GAC	TGT	GGC	CGG	CTG	1387
Met	Val	Glu	Asn	Gly	Arg	Phe	Ser	Gly	Phe	Ile	Asp	Cys	Gly	Arg	Leu	
			200					205					210			
GGT	GTG	GCG	GAC	CGC	TAT	CAG	GAC	ATA	GCG	TTG	GCT	ACC	CGT	GAT	ATT	1435
Gly	Val	Ala	Asp	Arg	Tyr	Gln	Asp	Ile	Ala	Leu	Ala	Thr	Arg	Asp	Ile	
		215					220					225				

GCT GAA GAG CTT GGC GGC GAA TGG GCT GAC CGC TTC CTC GTG CTT TAC Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe Leu Val Leu Tyr 230 235 240	1483
GGT ATC GCC GCT CCC GAT TCG CAG CGC ATC GCC TTC TAT CGC CTT CTT Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe Tyr Arg Leu Leu 245 250 255 260	1531
GAC GAG TTC TTC TGA GCGGGACTCT GGGGTTCGAA ATGACCGACC AAGCGACGCC Asp Glu Phe Phe 264	1586
CA TTG AGG GCG CAA GAG GAG AAA TGG ATT GAC CAA GAG ATC GTG GCT Leu Arg Ala Gln Glu Glu Lys Trp Ile Asp Gln Glu Ile Val Ala 197 200 205 210	1633
GTT ACG GAT GAA CAG TTC GAT TTA GAG GGC TAC AAC AGT CGA GCA ATT Val Thr Asp Glu Gln Phe Asp Leu Glu Gly Tyr Asn Ser Arg Ala Ile 215 220 225	1681
GAA CTG CCT CGG AAG GCA AAA TTG TTG ATC GTG ACA GTC ATC CGC GGC Glu Leu Pro Arg Lys Ala Lys Leu Leu Ile Val Thr Val Ile Arg Gly 230 235 240	1729
CTA GCA GTC TTT GAA GCC CTT TCC CGA TTG AAG CCT GTT CAT TCT GGC Leu Ala Val Phe Glu Ala Leu Ser Arg Leu Lys Pro Val His Ser Gly 245 250 255	1777
GGG GTG CAG ACT GCG GGC AAC AGC TGT GCC GTA GTG GAC GGC GCC GCG Gly Val Gln Thr Ala Gly Asn Ser Cys Ala Val Val Asp Gly Ala Ala 260 265 270 275	1825
GCG GCT TTG GTG GCT CGA GAG TCG TCT GCG ACA CAG CCG GTC TTG GCT Ala Ala Leu Val Ala Arg Glu Ser Ser Ala Thr Gln Pro Val Leu Ala 280 285 290	1873
AGG ATA CTG GCT ACC TCC GTA GTC GGG ATC GAG CCC GAG CAT ATG GGG Arg Ile Leu Ala Thr Ser Val Val Gly Ile Glu Pro Glu His Met Gly 295 300 305	1921
CTC GGC CCT GCG CCC GCG ATT CGC CTG CTG CTT GCG CGT AGT GAT CTT Leu Gly Pro Ala Pro Ala Ile Arg Leu Leu Leu Ala Arg Ser Asp Leu 310 315 320	1969
AGT TTG AGG GAT ATC GAC CTC TTT GAG ATA AAC GAG GCG CAG GCC GCC Ser Leu Arg Asp Ile Asp Leu Phe Glu Ile Asn Glu Ala Gln Ala Ala 325 330 335	2017
CAA GTT CTA GCG GTA CAG CAT GAA TTG GGT ATT GAG CAC TCA AAA CTT Gln Val Leu Ala Val Gln His Glu Leu Gly Ile Glu His Ser Lys Leu 340 345 350 355	2065
AAT ATT TGG GGC GGG GCC ATT GCA CTT GGA CAC CCG CTT GCC GCG ACC Asn Ile Trp Gly Gly Ala Ile Ala Leu Gly His Pro Leu Ala Ala Thr 360 365 370	2113

GGA TTG CGT CTC TGC ATG ACC CTC GCT CAC CAA TTG CAA GCT AAT AAC	2161
Gly Leu Arg Leu Cys Met Thr Leu Ala His Gln Leu Gln Ala Asn Asn	
375 380 385	
TTT CGA TAT GGA ATT GCC TCG GCA TGC ATT GGT GGG GGA CAG GGG ATG	2209
Phe Arg Tyr Gly Ile Ala Ser Ala Cys Ile Gly Gly Gly Gln Gly Met	
390 395 400	
GCG GTT CTT TTA GAG AAT CCC CAC TTC GGT TCG TCC TCT GCA CGA AGT	2257
Ala Val Leu Leu Glu Asn Pro His Phe Gly Ser Ser Ser Ala Arg Ser	
405 410 415	
TCG ATG ATT AAC AGA GTT GAC CAC TAT CCA CTG AGC TAA CGGGCATCTC	2306
Ser Met Ile Asn Arg Val Asp His Tyr Pro Leu Ser	
420 425 430 431	
CTTTGTTGCT TTGAGGTGGC GCACGAAGGA GGGCTCGAAA ATCTCTGCTA AAAACAAGAA	2366
GAAGGAACAG GGAACATGAT TAGTTTCGCT CGTATGGCAG AAAGTTTAGG AGTCCAGGCT	2426
AAACTTGCCC TTGCCTTCGC ACTCGTATTA TGTGTCGGGC TGATTGTTAC CGGCACGGGT	2486
TTCTACAGTG TACATACCTT GTCAGGGTTG GTGGGAATTC	2526

FIG. 2p:

GAATTCGCG GTCGGCGAAA GTTGATGCGC TGTATCGTGG TGAAGATCAA TCCATGCTGC	60
GTGACGAGGC CACACT GTG AGT TGG TCA GGG GGG GCT TAC TCG GCG TTT TCC	112
Met Ser Trp Ser Gly Gly Ala Tyr Ser Ala Phe Ser	
1 5 10	
GAC ACT GCG TTG GTT GCG GCA GTG CGC ACC CCC TGG ATT GAT TGC GGG	160
Asp Thr Ala Leu Val Ala Ala Val Arg Thr Pro Trp Ile Asp Cys Gly	
15 20 25	
GGT GCC CTG TCG CTG GTG TCG CCT ATC GAC TTA GGG GTA AAG GTC GCT	208
Gly Ala Leu Ser Leu Val Ser Pro Ile Asp Leu Gly Val Lys Val Ala	
30 35 40	
CGC CAA GTT CTG ATC CCT GCG TCG GTT GAA CCA CAA ATG GTC GAT AGC	256
Arg Glu Val Leu Met Arg Ala Ser Leu Glu Pro Gln Met Val Asp Ser	
45 50 55 60	
GTA CTC GCA GGC TCT ATG GCT CAA GCA AGC TTT GAT GCT TAC CTG CTC	304
Val Leu Ala Gly Ser Met Ala Gln Ala Ser Phe Asp Ala Tyr Leu Leu	
65 70 75	
CCG CGG CAC ATT GGC TTG TAC AGC GGT GTT CCC AAG TCG GTT CCG GCC	352
Pro Arg His Ile Gly Leu Tyr Ser Gly Val Pro Lys Ser Val Pro Ala	
80 85 90	
TTG GGG GTG CAG CGC ATT TGC GGC ACA GGC TTC GAA CTG CTT CGG CAG	400
Leu Gly Val Gln Arg Ile Cys Gly Thr Gly Phe Glu Leu Leu Arg Gln	
95 100 105	
GCC GGC GAG CAG ATT TCC CAA GGC GCT GAT CAC GTG CTG TGT GTC GCG	448
Ala Gly Glu Gln Ile Ser Gln Gly Ala Asp His Val Leu Cys Val Ala	
110 115 120	
GCA GAG TCC ATG TCG CGT AAC CCC ATC GCG TCG TAT ACA CAC CGG GGC	496
Ala Glu Ser Met Ser Arg Asn Pro Ile Ala Ser Tyr Thr His Arg Gly	
125 130 135 140	
GGG TTC CGC CTC GGT GCG CCC GTT GAG TTC AAG GAT TTT TTG TGG GAG	544
Gly Phe Arg Leu Gly Ala Pro Val Glu Phe Lys Asp Phe Leu Trp Glu	
145 150 155	
GCA TTG TTT GAT CCT GCT CCA GGA CTC GAC ATG ATC GCT ACC GCA GAA	592
Ala Leu Phe Asp Pro Ala Pro Gly Leu Asp Met Ile Ala Thr Ala Glu	
160 165 170	
AAC CTG GGGGAGAGGC GGTTTGCGTA TTGGGCGCAT GCATAAAAAC TGTTGTAATT	648
Asn Leu	
174	
CATTAAGCAT TCTGCCGACA TGGAAGCCAT CACAAACGGC ATGATGAACC TGAATCGCCA	708
GCGGCATCAG CACCTTGTCG CCTTGCGTAT AATATTTGCC CATGGACGCA CACCGTGGAA	768
ACGGATGAAG GCACGAAECC AGTTGACATA AGCCTGTTCG GTTCGTAAAC TGTAATGCAA	828
GTAGCGTATG CGCTCACGCA ACTGGTCCAG AACCTTGACC GAACGCAGCG GTGGTAACGG	888

CGCAGTGGCG	GTTTT	CATGG	CTTGTT	TATGA	CTGTTTTTTTT	GTACAGTCTA	TGCCTCGGGC	948								
ATCCAAGC	AGCAAGCGCG	TTACGCCGTG	GGTCGATGTTG	ATGTTATGGA	GCAGCAACG	1007										
ATG	TTA	CGC	AGC	AGC	AAC	GAT	GTT	ACG	CAG	CAG	GGC	AGT	CGC	CCT	AAA	1055
Met	Leu	Arg	Ser	Ser	Asn	Asp	Val	Thr	Gln	Gln	Gly	Ser	Arg	Pro	Lys	
1				5					10					15		
ACA	AAG	TTA	GGT	GGC	TCA	AGT	ATG	GGC	ATC	ATT	CGC	ACA	TGT	AGG	CTC	1103
Thr	Lys	Leu	Gly	Gly	Ser	Ser	Met	Gly	Ile	Ile	Arg	Thr	Cys	Arg	Leu	
			20					25					30			
GGC	CCT	GAC	CAA	GTC	AAA	TCC	ATG	CGG	GCT	GCT	CTT	GAT	CTT	TTC	GGT	1151
Gly	Pro	Asp	Gln	Val	Lys	Ser	Met	Arg	Ala	Ala	Leu	Asp	Leu	Phe	Gly	
		35					40					45				
CGT	GAG	TTC	GGA	GAC	GTA	GCC	ACC	TAC	TCC	CAA	CAT	CAG	CCG	GAC	TCC	1199
Arg	Glu	Phe	Gly	Asp	Val	Ala	Thr	Tyr	Ser	Gln	His	Gln	Pro	Asp	Ser	
	50					55					60					
GAT	TAC	CTC	GGG	AAC	TTG	CTC	CGT	AGT	AAG	ACA	TTC	ATC	GCG	CTT	GCT	1247
Asp	Tyr	Leu	Gly	Asn	Leu	Leu	Arg	Ser	Lys	Thr	Phe	Ile	Ala	Leu	Ala	
65					70					75				80		
GCC	TTC	GAC	CAA	GAA	GCG	GTT	GTT	GGC	GCT	CTC	GCG	GCT	TAC	GTT	CTG	1295
Ala	Phe	Asp	Gln	Glu	Ala	Val	Val	Gly	Ala	Leu	Ala	Ala	Tyr	Val	Leu	
			85					90						95		
CCC	AGG	TTT	GAG	CAG	CCG	CGT	AGT	GAG	ATC	TAT	ATC	TAT	GAT	CTC	GCA	1343
Pro	Arg	Phe	Glu	Gln	Pro	Arg	Ser	Glu	Ile	Tyr	Ile	Tyr	Asp	Leu	Ala	
			100					105					110			
GTC	TCC	GGC	GAG	CAC	CGG	AGG	CAG	GGC	ATT	GCC	ACC	GCG	CTC	ATC	AAT	1391
Val	Ser	Gly	Glu	His	Arg	Arg	Gln	Gly	Ile	Ala	Thr	Ala	Leu	Ile	Asn	
		115					120					125				
CTC	CTC	AAG	CAT	GAG	GCC	AAC	GCG	CTT	GGT	GCT	TAT	GTG	ATC	TAC	GTG	1439
Leu	Leu	Lys	His	Glu	Ala	Asn	Ala	Leu	Gly	Ala	Tyr	Val	Ile	Tyr	Val	
		130				135					140					
CAA	GCA	GAT	TAC	GGT	GAC	GAT	CCC	GCA	GTG	GCT	CTC	TAT	ACA	AAG	TTG	1487
Gln	Ala	Asp	Tyr	Gly	Asp	Asp	Pro	Ala	Val	Ala	Leu	Tyr	Thr	Lys	Leu	
145					150				155					160		
GGC	ATA	CGG	GAA	GAA	GTG	ATG	CAC	TTT	GAT	ATC	GAC	CCA	AGT	ACC	GCC	1535
Gly	Ile	Arg	Glu	Glu	Val	Met	His	Phe	Asp	Ile	Asp	Pro	Ser	Thr	Ala	
			165					170						175		
ACC	TAA	CAATTCGTTT	AAGCCGAGAT	CGGCTTCCCA	TTG	AGG	GCG	CAA	GAG	GAG	1589					
Thr											Leu	Arg	Ala	Gln	Glu	Glu
177											197			200		
AAA	TGG	ATT	GAC	CAA	GAG	ATC	GTG	GCT	GTT	ACG	GAT	GAA	CAG	TTC	GAT	1637
Lys	Trp	Ile	Asp	Gln	Glu	Ile	Val	Ala	Val	Thr	Asp	Glu	Gln	Phe	Asp	
		205					210					215				

TTA GAG GGC TAC AAC AGT CGA GCA ATT GAA CTG CCT CGG AAG GCA AAA	1685
Leu Glu Gly Tyr Asn Ser Arg Ala Ile Glu Leu Pro Arg Lys Ala Lys	
220 225 230	
TTG TTG ATC GTG ACA GTC ATC CGC GGC CTA GCA GTC TTT GAA GCC CTT	1733
Leu Leu Ile Val Thr Val Ile Arg Gly Leu Ala Val Phe Glu Ala Leu	
235 240 245 250	
TCC CGA TTG AAG CCT GTT CAT TCT GGC GGG GTG CAG ACT GCG GGC AAC	1781
Ser Arg Leu Lys Pro Val His Ser Gly Gly Val Gln Thr Ala Gly Asn	
255 260 265	
AGC TGT GCC GTA GTG GAC GGC GCC GCG GCG GCT TTG GTG GCT CGA GAG	1829
Ser Cys Ala Val Val Asp Gly Ala Ala Ala Leu Val Ala Arg Glu	
270 275 280	
TCG TCT GCG ACA CAG CCG GTC TTG GCT AGG ATA CTG GCT ACC TCC GTA	1877
Ser Ser Ala Thr Gln Pro Val Leu Ala Arg Ile Leu Ala Thr Ser Val	
285 290 295	
GTC GGG ATC GAG CCC GAG CAT ATG GGG CTC GGC CCT GCG CCC GCG ATT	1925
Val Gly Ile Glu Pro Glu His Met Gly Leu Gly Pro Ala Pro Ala Ile	
300 305 310	
CGC CTG CTG CTT GCG CGT AGT GAT CTT AGT TTG AGG GAT ATC GAC CTC	1973
Arg Leu Leu Leu Ala Arg Ser Asp Leu Ser Leu Arg Asp Ile Asp Leu	
315 320 325 330	
TTT GAG ATA AAC GAG GCG CAG GCC GCC CAA GTT CTA GCG GTA CAG CAT	2021
Phe Glu Ile Asn Glu Ala Gln Ala Ala Val Leu Ala Val Gln His	
335 340 345	
GAA TTG GGT ATT GAG CAC TCA AAA CTT AAT ATT TGG GGC GGG GCC ATT	2069
Glu Leu Gly Ile Glu His Ser Lys Leu Asn Ile Trp Gly Gly Ala Ile	
350 355 360	
GCA CTT GGA CAC CCG CTT GCC GCG ACC GGA TTG CGT CTC TGC ATG ACC	2117
Ala Leu Gly His Pro Leu Ala Ala Thr Gly Leu Arg Leu Cys Met Thr	
365 370 375	
CTC GCT CAC CAA TTG CAA GCT AAT AAC TTT CGA TAT GGA ATT GCC TCG	2165
Leu Ala His Gln Leu Gln Ala Asn Asn Phe Arg Tyr Gly Ile Ala Ser	
380 385 390	
GCA TGC ATT GGT GGG GGA CAG GGG ATG GCG GTT CTT TTA GAG AAT CCC	2213
Ala Cys Ile Gly Gly Gly Gln Gly Met Ala Val Leu Leu Glu Asn Pro	
395 400 405 410	
CAC TTC GGT TCG TCC TCT GCA CGA AGT TCG ATG ATT AAC AGA GTT GAC	2261
His Phe Gly Ser Ser Ser Ala Arg Ser Ser Met Ile Asn Arg Val Asp	
415 420 425	
CAC TAT CCA CTG AGC TAA CGGGCATCTC CTTTGTGCT TTGAGGTGGC	2309
His Tyr Pro Leu Ser	
430-431	

GCACGAAGGA GGGCTCGAAA ATCTCTGCTA AAAACAAGAA GAAGGAACAG GGAACATGAT	2369
TAGTTTCGCT CGTATGGCAG AAAGTTTAGG AGTCCAGGCT AAAGTTGCCC TTGCCTTCGC	2429
ACTCGTATTA TGTGTCGGGC TGATTGTTAC CGGCACGGGT TTCTACAGTG TACATACCTT	2489
GTCAGGGTTG GTGGGAATTC	2509

FIG. 2q:

GAATTCGCG	GTCGGCGAAA	GTTGATGCGC	TGTATCGTGG	TGAAGATCAA	TCCATGCTGC	60
GTGACGAGGC	CACACT	GTG AGT	TGG TCA	GGG GGG	GCT TAC TCG GCG TTT TCC	112
		Met Ser Trp Ser	Gly Gly Ala Tyr	Ser Ala Phe Ser		
	1		5		10	
GAC ACT GCG TTG GTT GCG GCA GTG CGC ACC CCC TGG ATT GAT TGC GGG	15	20	25	208		
Asp Thr Ala Leu Val Ala Ala Val Arg Thr Pro Trp Ile Asp Cys Gly						
GGT GCC CTG TCG CTG GTG TCG CCT ATC GAC TTA GGG GTA AAG GTC GCT	30	35	40	208		
Gly Ala Leu Ser Leu Val Ser Pro Ile Asp Leu Gly Val Lys Val Ala						
CGC CAA CTT CTC ATC CCT CCC TCG CTT CAA CCA CAA ATC CTC GAT AGC	45	50	55	60	256	
Arg Glu Val Leu Met Arg Ala Ser Leu Glu Pro Gln Met Val Asp Ser						
GTA CTC GCA GGC TCT ATG GCT CAA GCA AGC TTT GAT GCT TAC CTG CTC	65	70	75	304		
Val Leu Ala Gly Ser Met Ala Gln Ala Ser Phe Asp Ala Tyr Leu Leu						
CCG CGG CAC ATT GGC TTG TAC AGC GGT GTT CCC AAG TCG GTT CCG GCC	80	85	90	352		
Pro Arg His Ile Gly Leu Tyr Ser Gly Val Pro Lys Ser Val Pro Ala						
TTG GGG GTG CAG CGC ATT TGC GGC ACA GGC TTC GAA CTG CTT CGG CAG	95	100	105	400		
Leu Gly Val Gln Arg Ile Cys Gly Thr Gly Phe Glu Leu Leu Arg Gln						
GCC GGC GAG CAG ATT TCC CAA GGC GCT GAT CAC GTG CTG TGT GTC GCG	110	115	120	448		
Ala Gly Glu Gln Ile Ser Gln Gly Ala Asp His Val Leu Cys Val Ala						
GCA GAG TCC ATG TCG CGT AAC CCC ATC GCG TCG TAT ACA CAC CGG GGC	125	130	135	496		
Ala Glu Ser Met Ser Arg Asn Pro Ile Ala Ser Tyr Thr His Arg Gly						
GGG TTC CGC CTC GGT GCG CCC GTT GAG TTC AAG GAT TTT TTG TGG GAG	145	150	155	544		
Gly Phe Arg Leu Gly Ala Pro Val Glu Phe Lys Asp Phe Leu Trp Glu						
GCA TTG TTT GAT CCT GCT CCA GGA CTC GAC ATG ATC GCT ACC GCA GAA	160	165	170	592		
Ala Leu Phe Asp Pro Ala Pro Gly Leu Asp Met Ile Ala Thr Ala Glu						
AAC CTG GCG CGC A TTG AGG GCG CAA GAG GAG AAA TGG ATT GAC CAA GAG	175	176	197	200	205	641
Asn Leu Ala Arg Leu Arg Ala Gln Glu Glu Lys Trp Ile Asp Gln Glu						
ATC GTG GCT GTT ACG GAT GAA CAG TTC GAT TTA GAG GGC TAC AAC AGT	210	215	220	689		
Ile Val Ala Val Thr Asp Glu Gln Phe Asp Leu Glu Gly Tyr Asn Ser						
CGA GCA ATT GAA-CTG-CCT CGG AAG GCA AAA TTG TTG ATC GTG ACA GTC	225	230	235	240	737	
Arg Ala Ile Glu Leu Pro Arg Lys Ala Lys Leu Leu Ile Val Thr Val						

ATC CGC GGC CTA GCA GTC TTT GAA GCC CTT TCC CGA TTG AAG CCT GTT	785
Ile Arg Gly Leu Ala Val Phe Glu Ala Leu Ser Arg Leu Lys Pro Val	
245 250 255	
CAT TCT GGC GGG GTG CAG ACT GCG GGC AAC AGC TGT GCC GTA GTG GAC	833
His Ser Gly Gly Val Gln Thr Ala Gly Asn Ser Cys Ala Val Val Asp	
260 265 270	
GGC GCC GCG GCG GCT TTG GTG GCT CGA GAG TCG TCT GCG ACA CAG CCG	881
Gly Ala Ala Ala Leu Val Ala Arg Glu Ser Ser Ala Thr Gln Pro	
275 280 285	
GTC TTG GCT AGG ATA CTG GCT ACC TCC GTA GTC GGG ATC GAG CCC GAG	929
Val Leu Ala Arg Ile Leu Ala Thr Ser Val Val Gly Ile Glu Pro Glu	
290 295 300	
CAT ATG GGG CTC GGC CCT GCG CCC GCG ATT CGC CTG CTG CTT GCG CGT	977
His Met Gly Leu Gly Pro Ala Pro Ala Ile Arg Leu Leu Leu Ala Arg	
305 310 315 320	
AGT GAT CTT AGT TTG AGG GAT ATC GAC CTC TTT GAG ATA AAC GAG GCG	1025
Ser Asp Leu Ser Leu Arg Asp Ile Asp Leu Phe Glu Ile Asn Glu Ala	
325 330 335	
CAG GCC GCC CAA GTT CTA GCG GTA CAG CAT GAA TTG GGT ATT GAG CAC	1073
Gln Ala Ala Gln Val Leu Ala Val Gln His Glu Leu Gly Ile Glu His	
340 345 350	
TCA AAA CTT AAT ATT TGG GGC GGG GCC ATT GCA CTT GGA CAC CCG CTT	1121
Ser Lys Leu Asn Ile Trp Gly Gly Ala Ile Ala Leu Gly His Pro Leu	
355 360 365	
GCC GCG ACC GGA TTG CGT CTC TGC ATG ACC CTC GCT CAC CAA TTG CAA	1169
Ala Ala Thr Gly Leu Arg Leu Cys Met Thr Leu Ala His Gln Leu Gln	
370 375 380	
GCT AAT AAC TTT CGA TAT GGA ATT GCC TCG GCA TGC ATT GGT GGG GGA	1217
Ala Asn Asn Phe Arg Tyr Gly Ile Ala Ser Ala Cys Ile Gly Gly Gly	
385 390 395 400	
CAG GGG ATG GCG GTT CTT TTA GAG AAT CCC CAC TTC GGT TCG TCC TCT	1265
Gln Gly Met Ala Val Leu Leu Glu Asn Pro His Phe Gly Ser Ser Ser	
405 410 415	
GCA CGA AGT TCG ATG ATT AAC AGA GTT GAC CAC TAT CCA CTG AGC TAA	1313
Ala Arg Ser Ser Met Ile Asn Arg Val Asp His Tyr Pro Leu Ser	
420 425 430 431	
CGGGCATCTC CTTTGTGCT TTGAGGTGGC GCACGAAGGA GGGCTCGAAA ATCTCTGCTA	1373
AAAACAAGAA GAAGGAACAG GGAACATGAT TAGTTTCGCT CGTATGGCAG AAAGTTTAGG	1433
AGTCCAGGCT AAACCTGCCC TTGCCTTCGC ACTCGTATTA TGTGTCGGGC TGATTGTTAC	1493
CGGCACGGGT TTCTACAGTG TACATACCTT GTCAGGGTTG GTGGGAATTC	1543

FIG. 2r:

Sequenz 1

CTGCAGCCAG	GGCTGAAAAG	GAGGGATTCA	GTGAGGTCAT	GAAGGGAGGG	GACGGCGCCT	60
GGCTCCAATT	GCTCGATGGC	GCCGCGATTG	AGTGTCTTGG	GCGCGGTCTT	GGAGAGTTCG	120
GCTAGGGAGA	TAAATTTGCT	GGCCATGGTG	GCGGCCCTG	ATGGGTGGGA	TGATTTTCTG	180
CATTCTGCAT	CATGAAATTC	ATGAAATCAT	CACTTTTCGG	GGGGTGGGTG	CACGGGATTG	240
AAGTTTGCTA	GGAGAGTGCA	TTGCTCGTAA	GCCCAGGAAG	CACGCGGGTT	TCAGGATGGT	300
GCATGGAAAT	GGCATGAGCT	TTGCTGGATA	TGATTAGAGA	CATTAACAT	TTTGCGCGAA	360
TGGAAGCACG	ATTCTCGCC	CGGTAGAGCG	GTAACCGCGA	CATTACAGGAC	CGTAAAAAGG	420
AAAGAGCATG	CAACTGACCA	ACAAGAAAAT	CGTCGTCACC	GGAGTGTCTT	CCGGTATCGG	480
TGCCGAAACT	GCCCCGGTTC	TGCGCTCTCA	CGGCGCCACA	GTGATTGGCG	TAGATCGCAA	540
CATGCCGAGC	CTGACTCTGG	ATGCTTTCGT	TCAGGCTGAC	CTGAGCCATC	CTGAAGGCAT	600
CGATAAGGCC	ATCGGGACAG	CAAGCGAACC	GGAATTGCCA	GCTGGGGCGC	GCTCTGGTAA	660
GGTTGGGAAG	CCCTGCAAAG	TAAACTGGAT	GGCTTTCCTG	CCGCCAAGGA	TCTGATGGCG	720
CAGGGGATCA	AGATCTGATC	AAGAGACAGG	ATGAGGATCG	TTTCGCATGA	TTGAACAAGA	780
TGGATTGCAC	GCAGTTTCTC	CGGCCGCTTG	GGTGGAGAGG	CTATTCGGCT	ATGACTGGGC	840
ACAACAGACA	ATCGGCTGCT	CTGATGCCGC	CGTGTTCCGG	CTGTCAGCGC	AGGGGCGCCC	900
GGTTCTTTTT	GTCAAGACCG	ACCTGTCCGG	TGCCCTGAAT	GAATGACAGG	ACGAGGCAGC	960
GCGGCTATCG	TGGCTGGCCA	CGACGGGCGT	TCCTTGCGCA	GCTGTGCTCG	ACGTTGTAC	1020
TGAAGCGGGA	AGGGACTGGC	TGCTATTGGG	CGAAGTGCCG	GGGCAGGATC	TCCTGTATC	1080
TCACCTTGCT	CCTGCCGAGA	AAGTATCCAT	CATGGCTGAT	GCAATGCGGC	GGCTGCATAC	1140
GCTTGATCCG	GCTACCTGCC	CATTGACCA	CCAAGCGAAA	CATCGCATCG	AGCGAGCACG	1200
TACTCGGATG	GAAGCCGGTC	TTGTCGATCA	GGATGATCTG	GACGAAGAGC	ATCAGGGGCT	1260
CGCGCCAGCC	GAAGTGTTCG	CCAGGCTCAA	GGCGCGCATG	CCCACGCGCG	AGGATCTCGT	1320
CGTGACCCAT	GGCGATGCCCT	GCTTGCCGAA	TATCATGGTG	GAAAATGGCC	GCTTTTCTGG	1380
ATTCATCGAC	TGTGGCCGGC	TGGGTGTGGC	GGACCGCTAT	CAGGACATAG	CGTTGGCTAC	1440
CCGTGATATT	GCTGAAGAGC	TTGGCGGCGA	ATGGGCTGAC	CGCTTCCTCG	TGCTTTACGG	1500
TATCGCCGCT	CCCGATTTCG	AGCGCATCGC	CTTCTATCGC	CTTCTTGACG	AGTTCTTCTG	1560
AGCGGGACTC	TGGGGTTTCGA	AATGACCGAC	CAAGCGACGC	CCTGGCCGCG	GTGATTGCAT	1620
TCATGTGTGC	TGAGGAGTCA	CGTTGGATCA	ACGGCATAAA	TATCCAGTG	GACGGAGGTT	1680
TGGCATCGAC	CTACGTGTAA	GTTCGTGGAC	GCCCTTTGCA	CGCGCACTAT	ATCTCTATGC	1740
AGCAGCTGAA	AGCAGCTTTG	GTTTTGATCG	GAGGTAGCGG	GCGGAAAGGT	GCAGAATGTC	1800
TAAATAATAA	AGGATTCTTG	TGAAGCTTTA	GTTGTCCGTA	AACGAAAATA	AAAATAAAGA	1860
GGAATGATAT	GAAAGCAAGT	AGATCAGTCT	GCACTTTCAA	AATAGCTACC	CTGGCAGGCG	1920
CCATTTATGC	AGCGCTGCCA	ATGTCAGCTG	CAAACCTCGAT	GCAGCTGGAT	GTAGGTAGCT	1980
CGGATTGGAC	GGTGCGTTGG	GGACAACACC	CTCAAGTATA	GCCTTGCCTC	TCGCCTGAAT	2040
GAGCAAGACT	CAAGTCTGAC	AAATGCGCCG	ACTGTCAATG	GTTATATCCG	GATATTCAAA	2100
GTCAGGGTGA	TCGTAACTTT	GACCGGGGGC	TTGGTATCCA	ATCGTCTCGA	TATTCTGGCT	2160
GCAG						2164

Sequenz 2

CTGCAGCCAG	GGCTGAAAAG	GAGGGATTCA	GTGAGGTCAT	GAAGGGAGGG	GACGGCGCCT	60
GGCTCCAATT	GCTCGATGGC	GCCGCGATTG	AGTGTCTTGG	GCGCGGTCTT	GGAGAGTTCTG	120
GCTAGGGAGA	TAAATTTGCT	GGCCATGGTG	GCGGGCCCCTG	ATGGGTTGGA	TGATTTTCTG	180
CATTCTGCAT	CATGAAATTC	ATGAAATCAT	CACTTTTTCGG	GGGGTGGGTG	CACGGGATTG	240
AAGGTTGCTA	GGAGAGTGCA	TTGCTCGTAA	GCCCAGGAAG	CACGCGGGTT	TCAGGATGGT	300
GCATGGAAAT	GGCATGAGCT	TTGCTGGATA	TGATTAGAGA	CATTAACAT	TTTGGCGGAA	360
TGGAAGCACG	ATTCTCTGCC	CGGTAGAGCG	GTAACCGCGA	CATTCAGGAC	CGTAAAAAGG	420
AAAGAGCATG	CAACTGACCA	ACAAGAAAAT	CGTCGTCACC	GGAGTGTCTT	CCGGTATCGG	480
TGCCGAAACT	GCCCCGCTTC	TGCGCTCTCA	CGGCGCCACA	GTGATTGGCG	TAGATCGCAA	540
CATGCCGAGC	CTGACTCTGG	ATGCTTTCGT	TCAGGCTGAC	CTGAGCCATC	CTGAGGGGAG	600
AGGCGGTTTG	CGTATTGGGC	GCATGCATAA	AAACTGTTGT	AATTCATTAA	GCATTCTGCC	660
GACATGGAAG	CCATCACAAA	CGGCATGATG	AACCTGAATC	GCCAGCGGCA	TCAGCACCTT	720
CTCGGCTTGC	GTATAATATT	TGCCCATGCA	CGCACACCGT	GGAAACGGAT	GAAGGCACGA	780
ACCCAGTTGA	CATAAGCCTG	TTCGGTTCGT	AAACTGTAAT	GCAAGTAGCG	TATGCGCTCA	840
CGCAACTGGT	CCAGAACCTT	GACCGAACGC	AGCGGTGGTA	ACGGCGCAGT	GGCGGTTTTT	900
ATGGCTTGTT	ATGACTGTTT	TTTTGTACAG	TCTATGCCTC	GGGCATCCAA	GCAGCAAGCG	960
CGTTACGCCG	TGGGTCGATG	TTTGATGTTA	TGGAGCAGCA	ACGATGTTAC	GCAGCAGCAA	1020
CGATGTTACG	CAGCAGGGCA	GTGCCCCTAA	AACAAAGTTA	GGTGGCTCAA	GTATGGGCAT	1080
CATTGCGACA	TGTAGGCTCG	GCCCTGACCA	AGTCAAATCC	ATGCGGGCTG	CTCTTGATCT	1140
TTTCGGTCGT	GAGTTCGGAG	ACGTAGCCAC	CTACTCCCAA	CATCAGCCGG	ACTCCGATTA	1200
CCTCGGGAAC	TTGCTCCGTA	GTAAGACATT	CATCGCGCTT	GCTGCCTTCG	ACCAAGAAGC	1260
GGTTGTTGGC	GCTCTCGCGG	CTTACGTTCT	GCCCAGGTTT	GAGCAGCCGC	GTAGTGAGAT	1320
CTATATCTAT	GATCTCGCAG	TCTCCGGCGA	GCACCGGAGG	CAGGGCATTG	CCACCGCGCT	1380
CATCAATCTC	CTCAAGCATG	AGGCCAACGC	GCTTGGTGCT	TATGTGATCT	ACGTGCAAGC	1440
AGATTACGGT	GACGATCCCG	CAGTGGCTCT	CTATACAAAG	TTGGGCATAC	GGGAAGAAGT	1500
GATGCACTTT	GATATCGACC	CAAGTACCGC	CACCTAACAA	TTCGTTCAAG	CCGAGATCGG	1560
CTTCCCTGAT	TGCATTATG	TGTGCTGAGG	AGTCACGTTG	GATCAACGGC	ATAAATATTC	1620
CAGTGGACGG	AGGTTTGGCA	TCGACCTACG	TGTAAGTTCTG	TGGACGCCCT	TTGCACGCGC	1680
ACTATATCTC	TATGCAGCAG	CTGAAAGCAG	CTTTGGTTTT	GATCGGAGGT	AGCGGGCGGA	1740
AAGGTGCAGA	ATGCTCTAAAT	AATAAAGGAT	TCTTGTGAAG	CTTTAGTTGT	CCGTAAACGA	1800
AAATAAAAAAT	AAAGAGGAAT	GATATGAAAG	CAAGTAGATC	AGTCTGCACT	TTCAAAATAG	1860
CTACCCTGGC	AGGCGCCATT	TATGCAGCGC	TGCCAATGTC	AGCTGCAAAC	TCGATGCAGC	1920
TGGATGTAGG	TAGCTCGGAT	TGGACGGTGC	GTGGGGGACA	ACACCCTCAA	GTATAGCCTT	1980
GCCTCTCGCC	TGAATGAGCA	AGACTCAAGT	CTGACAAATG	CGCCGACTGT	CAATGGTTAT	2040
ATCCGGATAT	TCAAAGTCAG	GGTGATCGTA	ACTTTGACCG	GGGGCTTGGT	ATCCAATCGT	2100
CTCGATATTC	TGGCTGCAG					2119

Sequenz 3

CTGCAGCCAG	GGCTGAAAAG	GAGGGATTCA	GTGAGGTCAT	GAAGGGAGGG	GACGGCGCCT	60
GGCTCCAATT	GCTCGATGGC	GCCGCGATTG	AGTGTCTTGG	GCGCGGTCTT	GGAGAGTTCG	120
GCTAGGGAGA	TAAATTTGCT	GGCCATGGTG	GCGGCCCCTG	ATGGGTGGGA	TGATTTTCTG	180
CATTCTGCAT	CATGAAATTC	ATGAAATCAT	CACTTTTCGG	GGGGTGGGTG	CACGGGATTG	240
AAGGTTGCTA	GGAGAGTGCA	TTGCTCGTAA	GCCCAGGAAG	CACGCGGGTT	TCAGGATGGT	300
GCATGGAAAT	GGCATGAGCT	TTGCTGGATA	TGATTAGAGA	CATTAACAT	TTTGGCGGAA	360
TGGAAGCACG	ATTCTTCGCC	CGGTAGAGCG	GTAACCGCGA	CATTCAGGAC	CGTAAAAAGG	420
AAAGAGCATG	CAACTGACCA	ACAAGAAAAAT	CGTCGTCACC	GGAGTGTCTT	CCGGTATCGG	480
TGCCGAAACT	GCCCGCGTTC	TGCGCTCTCA	GCGCGCCACA	GTGATTGGCG	TAGATCGCAA	540
CATGCCGAGC	CTGACTCTGG	ATGCTTTTCGT	TCAGGCTGAC	CTGAGCCATC	CTGAAGGCAT	600
CGATCAACGG	CATAAATATT	CCAGTGGACG	GAGGTTTGGC	ATCGACCTAC	GTGTAAGTTC	660
GTGGACGCCC	TTTGACGCG	CACTATATCT	CTATGCAGCA	GCTGAAAGCA	GCTTTGGTTT	720
TCATCCGACC	TACCGCCCCG	AAACSTCCAG	AATCTCTAAA	TAATAAAGCA	TTCTTCTCAA	780
GCTTTAGTTG	TCCGTAAACG	AAAATAAAAA	TAAAGAGGAA	TGATATGAAA	GCAAGTAGAT	840
CAGTCTGCAC	TTTCAAATA	GCTACCCTGG	CAGGCGCCAT	TTATGCAGCG	CTGCCAATGT	900
CAGCTGCAAA	CTCGATGCAG	CTGGATGTAG	GTAGCTCGGA	TTGGACGGTG	CGTTGGGGAC	960
AACACCCTCA	AGTATAGCCT	TGCCTCTCGC	CTGAATGAGC	AAGACTCAAG	TCTGACAAAT	1020
GCGCCGACTG	TCAATGGTTA	TATCCGGATA	TTCAAAGTCA	GGGTGATCGT	AACTTTGACC	1080
GGGGGCTTGG	TATCCAATCG	TCTCGATATT	CTGGCTGCAG			1120

Sequenz 4

GAATTCCGCG	TATCGCCCCG	TTCTATCAGC	GGGCCGCTTT	CGAAAGTCAT	GGTGTTAGCC	60
GGTAGGGTCT	TTTTCTTGGC	CATGCTTGTT	GCCTGAACCT	TCGTTGACAT	AGGGCAGAGG	120
TGCGTTTGCC	GCTTCGCTTC	GCGATGAACC	GCATCGAGAT	GCTGAGGTCA	GGATTTTTC	180
TTAACTCGCG	TAAGCATTCT	GTCATTTTTT	TGGTGGCTTT	GAACAGCCTG	ATGAAAGGTG	240
GTCTCGCCCT	TTGAGGCCGA	TTCTTGGGCG	CTTGGCGGCG	TCGAAGCGAT	GCTCCACTAC	300
CGATTAAGAT	AATTTAAATA	AGGAAACCGC	ATGGTTTCTT	ATGTGAATTT	GTCTGGCATA	360
CTCCAGCTCA	AGGGCAATTT	TTGGGCTATT	GGCTGAGCAG	TTGCCTCTAT	ATGGTTATTC	420
AGAATAACAA	TTGACTCCTC	AGGAGGTGAG	CGATGAGCAT	TCTTGGTTTG	AATGGTGCCC	480
CGGTGCGAGC	TGAGCAGCTG	GGCTCGGCTC	TTGATCGCAT	GAAGAAGGCG	CACCTGGAGC	540
AGGGGCTGTC	AAACTTGGAG	CTGCGTCTGA	GTAGGCTGGA	TCGTGCGATT	GCAATGCTTC	600
TGGAAATCG	TGAAGCAATT	GCCGACGCGG	TTTCTGCTGA	CTTTGGCAAT	CGCAGCCGTG	660
AGCAAACACT	GCTTTGCGAC	ATTGCTGGCT	CGGTGGCAAG	CCTGAAGGAT	AGCCGCGAGC	720
ACGTGGCCAA	ATGGATGGAG	CCCGAACATC	ACBAGGGGAT	GTTTCCAGCG	GCGGAGGCAG	780
GCGTTGAGTT	TCAGCCGCTG	GGTGTGCTTG	GGGTCAATAG	TCCCTGGAAC	TTCCCTATCG	840
TACTGGCCTT	TGGGCCGCTG	GCCGGCATAT	TCGCAGCAGG	TAATCGCGCC	ATGCTCAAGC	900
CGTCCGAGCT	TACCCCGCGG	ACTTCTGCCC	TGCTTGCGGA	GCTAATTGCT	CGTTACTTCG	960
ATGAAACTGA	GCTGACTACA	GTGCTGGGCG	ACGCTGAAGT	CGGTGCGCTG	TTCAGTGCTC	1020
AGCCTTTCGA	TCATCTGATC	TTCACCGGCG	GCACTGCCGT	GGCCAAGCAC	ATCATGCGTG	1080
CCGCGGCGGA	TAACCTAGTG	CCCGTTACCC	TGGAATTGGG	TGGCAAATCG	CCGGTGATCG	1140
TTTCCCGCAG	TGCAGATATG	GCGGACGTTG	CACAACGGGT	GTTGACGGTG	AAAACCTTCA	1200
ATGCCGGGCA	AATCTGTCTG	GCACCGGACT	ATGTGCTGCT	GCCGGAAGGG	ACAGCAAGCG	1260
AACCGGAATT	GCCAGCTGGG	GCGCCCTCTG	GTAAGGTTGG	GAAGCCCTGC	AAAGTAAACT	1320
GGATGGCTTT	CTTGCCGCCA	AGGATCTGAT	GGCGCAGGGG	ATCAAGATCT	GATCAAGAGA	1380
CAGGATGAGG	ATCGTTTTCG	ATGATTGAAC	AAGATGGATT	GCACGCAGGT	TCTCCGGCCG	1440
CTTGGGTGGA	GAGGCTATTC	GGCTATGACT	GGGCACAACA	GACAATCGGC	TGCTCTGATG	1500
CCGCGGTGTT	CCGGCTGTCA	GCGCAGGGGC	GCCCGGTTCT	TTTTGTCAAG	ACCGACCTGT	1560
CCGGTGCCCT	GAATGAACTG	CAGGACGAGG	CAGCGCGGCT	ATCGTGGCTG	GCCACGACGG	1620
GCGTTCCTTG	CGCAGCTGTG	CTCGACGTTG	TCACTGAAGC	GGGAAGGGAC	TGGCTGCTAT	1680
TGGGCGAAGT	GCCGGGGCAG	GATCTCCTGT	CATCTCACCT	TGCTCCTGCC	GAGAAAGTAT	1740
CCATCATGGC	TGATGCAATG	CGCGGGCTGC	ATACGCTTGA	TCCGGCTACC	TGCCCATTTC	1800
ACCACCAAGC	GAAACATCGC	ATCGAGCGAG	CACGTACTCG	GATGGAAGCC	GGTCTTGTCG	1860
ATCAGGATGA	TCTGGACGAA	GAGCATCAGG	GGCTCGCGCC	AGCCGAAGTG	TTCCGCCAGC	1920
TCAAGGCGCG	CATGCCCCGAC	GGCGAGGATC	TCGTGCTGAC	CCATGGCGAT	GCCTGCTTGC	1980
CGAATATCAT	GGTGAAAAAT	GGCCGCTTTT	CTGGATTTCAT	CGACTGTGGC	CGGCTGGGTG	2040
TGGCGGACCG	CTATCAGGAC	ATAGCGTTGG	CTACCCGTGA	TATTGCTGAA	GAGCTTGGCG	2100
GCGAATGGGC	TGACCGCTTC	CTCGTGCTTT	ACGGTATCGC	CGCTCCCGAT	TCGCAGCGCA	2160

TCGCCTTCTA	TCGCCTTCTT	GACGAGTTCT	TCTGAGCGGG	ACTCTGGGGT	TCGAAATGAC	2220
CGACCAAGCG	ACGCCC GCCA	TGCCAAGCCT	GTTCTCGTGC	AAAGTCCTGT	GGGTGAGTCG	2280
AACTTGCGA	TGCGCGCACC	CTACGGAGAA	GCGATCCACG	GACTGCTCTC	TGTCCTCCTT	2340
TCAACGGAGT	GTTAGAACCG	TTGGTAGTGG	TTTTGGACGG	GCCCAGGAGC	ATGCGCTTCT	2400
GGGCCCGTTT	CTTGAGTATT	CATTGGATAG	TCACGCGTGG	TAGCTTCGAG	CCTGCACAGC	2460
TGATGAGCAC	CCTGGAAGGC	GCGCTGTACG	CGGACGACTG	GGTTCATCTT	CGCCATTCAT	2520
GACGGAACCT	CGTTCCCCAG	TACCGCGATG	ACTATTTTGC	CTCTTCCGAT	GTCCGATTCC	2580
ACGCCGCCTG	ACGCTAAGCG	GGGGCGGGGG	CGCCCGCATC	CCAGCCCAGA	CAGCAACAAA	2640
TGAGTAGGCT	CTTGATGCC	GCGGCGGCTG	AGATTGGTAA	CGGCAATTTC	GTCAATGTGA	2700
CGATGGATTC	GATTGCCCGT	GCTGCCGGCG	TCTCAAAAAA	AACGCTGTAC	GTCTTGGTGG	2760
CGAGCAAGGA	AGAACTCATT	TCCCGGTTAG	TGGCTCGAGA	CATGTCCAAC	CTTGAGGAAT	2820
TC						2822

Sequenz 5

GAATTCGCG	TATCGCCCGG	TTCTATCAGC	GGGCCGCTTT	CGAAAGTCAT	GGTGTTAGCC	60
GGTAGGGTCT	TTTTCTTGGC	CATGCTTGTT	GCCTGAACCT	TCGTTGACAT	AGGGCAGAGG	120
TGCGTTTGCC	GCTTCGCTTC	GCGATGAACC	GCATCGAGAT	GCTGAGGTCA	GGATTTTCC	180
TTAACTCGCG	TAAGCATTCT	GTCATTTTTT	TGGTGGCTTT	GAACAGCCTG	ATGAAAGGTG	240
GTCTCGCCCT	TTGAGGCCGA	TTCTTGGGCG	CTTGGCGGCG	TCGAAGCGAT	GCTCCACTAC	300
CGATTAAGAT	AATTAAAAATA	AGGAAACCGC	ATGGTTTCTT	ATGTGAATTT	GTCTGGCATA	360
CTCCAGCTCA	AGGGCAATTT	TTGGGCTATT	GGCTGAGCAG	TTGCCTCTAT	ATGGTTATTC	420
AGAATAACAA	TTGACTCCTC	AGGAGGTCAG	CGATGAGCAT	TCTTGGTTTG	AATGGTGCCC	480
CGGTCGGAGC	TGAGCAGCTG	GGCTCGGCTC	TTGATCGCAT	GAAGAAGGCG	CACCTGGAGC	540
AGGGGCCCTG	AAACTTGGAG	CTGCGTCTGA	GTAGGCTGGA	TCGTGCGATT	GCAATGCTTC	600
TGGAAAATCG	TGAAGCAATT	GCCGACGCGG	TTTCTGCTGA	CTTTGGCAAT	CGCAGCCGTG	660
AGCAAACAC	GCTTTGCGAC	ATTGCTGGCT	CGGTGGCAAG	CCTGAAGGAT	AGCCGCGAGC	720
ACGTGGCCAA	ATGGATGGAG	CCCGAACATC	ACAAGGCGAT	GTTTCCAGGG	GCGGAGGGCAG	780
GCGTTGAGTT	TCAGCCGCTG	GGTGTCGTTG	GGGTCATTAG	TCCCTGGAAC	TTCCCTATCG	840
TACTGGCCTT	TGGGCCGCTG	GCCGGCATAT	TCGCAGCAGG	TAATCGCGCC	ATGCTCAAGC	900
CGTCCGAGCT	TACCCCGCGG	ACTTCTGCCC	TGCTTGCGGA	GCTAATTGCT	CGTTACTTCG	960
ATGAAACTGA	GCTGACTACA	GTGCTGGGCG	ACGCTGAAGT	CGGTGCGCTG	TTCAGTGCTC	1020
AGCCTTTTCA	TCATCTGATC	TTACCCGGCG	GCACTGCCGT	GGCCAAGCAC	ATCATGCGTG	1080
CCGCGGCGGA	TAACCTAGTG	CCCCTTACCC	TGGAATTGGG	TGGCAAATCG	CCGGTGATCG	1140
TTTCCCGCAG	TGCAGATATG	GCGGACGTTG	CACAACGGGT	GTTGACGGTG	AAAACCTTCA	1200
ATGCCGGGCA	AATCTGTCTG	GCACCGGACT	ATGTGCTGGG	GGAGAGGCGG	TTTGCGTATT	1260
GGGCGCATGC	ATAAAAACTG	TTGTAATTCA	TTAAGCATTC	TGCCGACATG	GAAGCCATCA	1320
CAAACGGCAT	GATGAACCTG	AATCGCCAGC	GGCATCAGCA	CCTTGTCGCC	TTGCGTATAA	1380
TATTTGCCCA	TGGACGCACA	CCGTGGAAAC	GGATGAAGGC	ACGAACCCAG	TTGACATAAG	1440
CCTGTTTCGGT	TCGTAAACTG	TAATGCAAGT	AGCGTATGCG	CTCACGCAAC	TGGTCCAGAA	1500
CCTTGACCGA	ACGCAGCGGT	GGTAACGGCG	CAGTGGCGGT	TTTCATGGCT	TGTTATGACT	1560
GTTTTTTTTGT	ACAGTCTATG	CCTCGGGCAT	CCAAGCAGCA	AGCGCGTTAC	GCCGTGGGTC	1620
GATGTTTGAT	GTTATGGAGC	AGCAACGATG	TTACGCAGCA	GCAACGATGT	TACGCAGCAG	1680
GGCAGTCGCC	CTAAAACAAA	GTTAGGTGGC	TCAAGTATGG	GCATCATTCG	CACATGTAGG	1740
CTCGGCCCTG	ACCAAGTCAA	ATCCATGCGG	GCTGCTCTTG	ATCTTTTCGG	TCGTGAGTTC	1800
GGAGACGTAG	CCACCTACTC	CCAACATCAG	CCGGACTCCG	ATTACCTCGG	GAACCTTGCTC	1860
CGTAGTAAGA	CATTCATCGC	GCTTGCTGCC	TTCGACCAAG	AAGCGGTTGT	TGGCGCTCTC	1920
GCGGCTTACG	TTCTGCCCAG	GTTTGAGCAG	CCGCGTAGTG	AGATCTATAT	CTATGATCTC	1980
GCAGTCTCCG	GCGAGCACCG	GAGGCAGGGC	ATTGCCACCG	CGCTCATCAA	TCTCCTCAAG	2040
CATGAGGCCA	ACGCGCTTGG	TGCTTATGTG	ATCTACGTGC	AAGCAGATTA	CGGTGACGAT	2100
CCCGCAGTGG	CTCTCTATAC	AAAGTTGGGC	ATACGGGAAG	AAGTGATGCA	CTTTGATATC	2160

GACCCAAGTA	CCGCCACCTA	ACAATTGTT	CAAGCCGAGA	TCGGCTTCCC	TGCAAAGTCC	2220
TGTGGGTGAG	TCGAACTGG	CGATGCGCGC	ACCCTACGGA	GAAGCGATCC	ACGGACTGCT	2280
CTCTGTCCTC	CTTTCAACGG	AGTGTTAGAA	CCGTTGGTAG	TGGTTTTGGA	CGGGCCCAGG	2340
AGCATGCGCT	TCTGGGCCCC	TTTCTTGAGT	ATTCATTGGA	TAGTCACGCG	TGGTAGCTTC	2400
GAGCCTGCAC	AGCTGATGAG	CACCCTGGAA	GGCGCGCTGT	ACGCGGACGA	CTGGGTTCAT	2460
CTTCGCCATT	CATGACGGAA	CTCCGTTCCC	CAGTACCGCG	ATGACTATTT	TGCCTCTTCC	2520
GATGTCCGAT	TCCACGCCGC	CTGACGCTAA	GCGGGGGCGG	GGGCGCCCGC	ATCCCAGCCC	2580
AGACAGCAAC	AAATGAGTAG	GCTCTTGAT	GCCGCGGCGG	CTGAGATTGG	TAACGGCAAT	2640
TTCGTCAATG	TGACGATGGA	TTCGATTGCC	CGTGCTGCCG	GCGTCTCAA	AAAAACGCTG	2700
TACGTCTTGG	TGGCGAGCAA	GGAAGAACTC	ATTTCCCGGT	TAGTGGCTCG	AGACATGTCC	2760
AACCTTGAGG	AATTC					2775

Sequenz 6

GAATTCGCG	TATCGCCCG	TTCTATCAGC	GGGCCGCTTT	CGAAAGTCAT	GGTGTTAGCC	60
GGTAGGGTCT	TTTTCTTGGC	CATGCTTGTT	GCCTGAACCT	TCGTTGACAT	AGGGCAGAGG	120
TGCGTTTGCC	GCTTCGCTTC	GCGATGAACC	GCATCGAGAT	GCTGAGGTCA	GGATTTTTCC	180
TTAACTCGCG	TAAGCATCT	GTCATTTTTT	TGGTGGCTTT	GAACAGCCTG	ATGAAAGGTG	240
GTCTCGCCCT	TTGAGGCCGA	TTCTTGGGCG	CTTGGCGGCG	TCGAAGCGAT	GCTCCACTAC	300
CGATTAAGAT	AATTAATAA	AGGAAACCGC	ATGGTTTCTT	ATGTGAATTT	GTCTGGCATA	360
CTCCAGCTCA	AGGGCAATTT	TTGGGCTATT	GGCTGAGCAG	TTGCCTCTAT	ATGGTTATTC	420
AGAATAACAA	TTGACTCCTC	AGGAGGTCAG	CGATGAGCAT	TCTTGGTTTG	AATGGTGCCC	480
CGGTGCGAGC	TGAGCAGCTG	GGCTCGGCTC	TTGATCGCAT	GAAGAAGGCG	CACCTGGAGC	540
AGGGGCGCTG	AACTTGGAG	CTGCGTCTGA	GTAGGCTGGA	TCGTGCGATT	GCAATGCCTC	600
TGGAAAATCG	TGAAGCAATT	GCCGACGCGG	TTTCTGCTGA	CTTGGCAAT	CGCAGCCGTG	660
AGCAAACACT	GCTTTGCGAC	ATTGCTGGCT	CGGTGGCAAG	CCTGAAGGAT	AGCCGCGAGC	720
ACGTGGCCAA	ATGATCGGCG	CCCCAACATC	ACAGGGCGAT	CTTTCAGCGG	CCCCAGCCAG	780
GCGTTGAGTT	TCAGCCGCTG	GGTGTGTTG	GGGTCATTAG	TCCCTGGAAC	TTCCCTATCG	840
TACTGGCCTT	TGGGCCGCTG	GCCGGCATAT	TCGCAGCAGG	TAATCGCGCC	ATGCTCAAGC	900
CGTCCGAGCT	TACCCCGCGG	ACTTCTGCCC	TGCTTGCGGA	GCTAATTGCT	CGTTACTTCG	960
ATGAAACTGA	GCTGACTACA	GTGCTGGGCG	ACGCTGAAGT	CGGTGCGCTG	TTCAGTGCTC	1020
AGCCTTTTCA	TCATCTGATC	TTCACCGGCG	GCACTGCCGT	GGCCAAGCAC	ATCATGCGTG	1080
CCGCGGCGGA	TAACCTAGTG	CCCGTTACCC	TGGAATTGGG	TGGCAAATCG	CCGGTGATCG	1140
TTTCCCGCAG	TGCAGATATG	GCGGACGTTG	CACAACGGGT	GTGACGGTG	AAAACCTTCA	1200
ATGCCGGGCA	AATCTGTCTG	GCACCGTGGG	TGAGTCGAAC	TTGGCGATGC	GCGCACCTTA	1260
CGGAGAAGCG	ATCCACGGAC	TGCTCTCTGT	CCTCCTTTCA	ACGGAGTGTT	AGAACCGTTG	1320
GTAGTGGTTT	TGGACGGGCC	CAGGAGCATG	CGCTTCTGGG	CCCGTTTCTT	GAGTATTTCAT	1380
TGGATAGTCA	CGCGTGCTAG	CTTCGAGCCT	GCACAGCTGA	TGAGCACCTT	GGAAGGCGCG	1440
CTGTACGCGG	ACGACTGGGT	TCATCTTCGC	CATTTCATGAC	GGAACCTCCG	TCCCCAGTAC	1500
CGCGATGACT	ATTTTGCCCTC	TTCCGATGTC	CGATTCCACG	CCGCTGACG	CTAAGCGGGG	1560
GCGGGGGCGC	CCGCATCCCA	GCCCAGACAG	CAACAAATGA	GTAGGCTCTT	GGATGCCGCG	1620
GCGGCTGAGA	TTGGTAACGG	CAATTTTCGT	AATGTGACGA	TGGATTTCAT	TGCCCCGTGCT	1680
GCCGGCGTCT	CAAAAAAAC	GCTGTACGTC	TTGGTGGCGA	GCAAGGAAGA	ACTCATTTC	1740
CGGTTAGTGG	CTCGAGACAT	GTCCAACCTT	GAGGAATTC			1779

Sequenz 7

CTGCAGCCGA	GCATCGATTG	AGCACTTTAC	CCAGCTGCGC	TGGCTGACCA	TTCAGAATGG	60
CCCGCGGCAC	TATCCAATCT	AAATCGATCT	TCGGGCGCCG	CGGGCATCAT	GCCCCGCGCG	120
CTCGCCTCAT	TTCAATCTCT	AACTTGATAA	AAACAGAGCT	GTTCTCCGGT	CTTGGTGGAT	180
CAAGGCCAGT	CGCGGAGAGT	CTCGAAGAGG	AGAGTACAGT	GAACGCCGAG	TCCACATTGC	240
AACCGCAGGC	ATCATCATGC	TCTGCTCAGC	CACGCTACCG	CAGTGTGTCG	ATTGGTCATC	300
CTCCGGTTGA	GGTTACGCAA	GACGCTGGAG	GTATTGTCCG	GATGCGTTCT	CTCGAGGCGC	360
TTCTTCCCTT	CCCGGGTCGA	ATTCTTGAGC	GTCTCGAGCA	TTGGGCTAAG	ACCCGTCCAG	420
AACAAACCTG	CGTTGCTGCC	AGGGCGGCAA	ATGGGGAATG	GCGTCGTATC	AGCTACGCGG	480
AAATGTTCCA	CAACGTCCGC	GCCATCGCAC	AGAGCTTGCT	TCCTTACGGA	CTATCGGCAG	540
AGCGTCCGCT	GCTTATCGTC	TCTGGAAATG	ACCTGGAACA	TCTTCAGCTG	GCATTTGGGG	600
CTATGTATGC	GGGCATTCCC	TATTGCCCGG	TGTCTCCTGC	TTATTCACTG	CTGTGCAAG	660
ATTTGGCGAA	GCTGCGTCAC	ATCGTAGGTC	TTCTGCAACC	GGGACTGGTC	TTTGCTGCCG	720
ATGCGAGCAG	TTTCCAGGGG	ACAGCAAGCG	AACCGGAATT	GCCAGCTGGG	GCGCCCTCTG	780
GTAAGGTTGG	GAAGCCCTGC	AAAGTAAACT	GGATGGCTTT	CTTGCCGCCA	AGGATCTGAT	840
GGCGCAGGGG	ATCAAGATCT	GATCAAGAGA	CAGGATGAGG	ATCGTTTCGC	ATGATTGAAC	900
AAGATGGATT	GCACGCAGGT	TCTCCGGCCG	CTTGGGTGGA	GAGGCTATTC	GGCTATGACT	960
GGGCACAACA	GACAATCGGC	TGCTCTGATG	CCGCCGTGTT	CCGGCTGTCA	GCGCAGGGGC	1020
GCCCCGTTCT	TTTTGTCAAG	ACCGACCTGT	CCGGTGCCCT	GAATGAACTG	CAGGACGAGG	1080
CAGCGCGGCT	ATCGTGGCTG	GCCACGACGG	GCGTTCCTTG	CGCAGCTGTG	CTCGACGTTG	1140
TCACTGAAGC	GGGAAGGGAC	TGGCTGCTAT	TGGGCGAAGT	GCCGGGGCAG	GATCTCCTGT	1200
CATCTCACCT	TGCTCCTGCC	GAGAAAGTAT	CCATCATGGC	TGATGCAATG	CGGCGGCTGC	1260
ATACGCTTGA	TCCGGCTACC	TGCCATTTCG	ACCACCAAGC	GAAACATCGC	ATCGAGCGAG	1320
CACGTACTCG	GATGGAAGCC	GGTCTTGTCG	ATCAGGATGA	TCTGGACGAA	GAGCATCAGG	1380
GGCTCGCGCC	AGCCGAACTG	TTCGCCAGGC	TCAAGGCGCG	CATGCCCCGAC	GGCGAGGATC	1440
TCGTCTGTAC	CCATGGCGAT	GCCTGCTTGC	CGAATATCAT	GGTGGAATAA	GGCCGCTTTT	1500
CTGGATTTCAT	CGACTGTGGC	CGGCTGGGTG	TGGCGGACCG	CTATCAGGAC	ATAGCGTTGG	1560
CTACCCGTGA	TATTGCTGAA	GAGCTTGCGG	GCGAATGGGC	TGACCGCTTC	CTCGTGCTTT	1620
ACGGTATCGC	CGCTCCCGAT	TCGCAGCGCA	TCGCCTTCTA	TCGCCTTCTT	GACGAGTTCT	1680
TCTGAGCGGG	ACTCTGGGGT	TCGAAATGAC	CGACCAAGCG	ACGCCCCGTG	TTTGCAATGG	1740
CGGTGCGCGA	AAGTTGATGC	GCTGTATCGT	GGTGAAGATC	AATCCATGCT	GCGTGACGAG	1800
GCCACACTGT	GAGTTGGTCA	GGGGGGGCTT	ACTCGGCGTT	TTCCGACACT	GCGTTGGTTG	1860
CGGCAGTGCG	CACCCCTGG	ATTGATTGCG	GGGGTGCCCT	GTCGCTGGTG	TCGCCTATCG	1920
ACTTAGGGGT	AAAGGTCGCT	CGCGAAGTTC	TGATGCGTGC	GTCGCTTGAA	CCACAAATGG	1980
TCGATAGCGT	ACTCGCAGGC	TCTATGGCTC	AAGCAAGCTT	TGATGCTTAC	CTGCTCCCGC	2040
GGCACATTGG	CTGTACAGC	GGTGTTCCTA	AGTCGGTTCC	GGCCTTGGGG	GTGCAGCGCA	2100
TTTGCGGCAC	AGGCTTCGAA	CTGCTTCGGC	AGGCCGGCGA	GCAGATTTC	CAAGGCGCTG	2160
ATCACGTGCT	GTGTGTCGCG	GGCTGCAG				2188

Sequenz 8

CTGCAGCCGA	GCATCGATTG	AGCACTTTAC	CCAGCTGCGC	TGGCTGACCA	TTCAGAATGG	60
CCC GCGGCAC	TATCCAATCT	AAATCGATCT	TCGGGCGCCG	CGGGCATCAT	GCCC GCGGCG	120
CTCGCCTCAT	TTCAATCTCT	AAC TTGATAA	AAACAGAGCT	GTTCTCCGGT	CTTGGTG GAT	180
CAAGGCCAGT	CGCGGAGAGT	CTCGAAGAGG	AGAGTACAGT	GAACGCCGAG	TCCACATTGC	240
AACCCAGGC	ATCATCATGC	TCTGCTCAGC	CACGCTACCG	CAGTGTGTCG	ATTGGTCATC	300
CTCCGGTTGA	GGTTACGCAA	GACGCTGGAG	GTATTGTCCG	GATGCGTTCT	CTCGAGGCGC	360
TTCTTCCCTT	CCC GGGTCGA	ATTCTTGAGC	GTCTCGAGCA	TTGGGCTAAG	ACCCGTCCAG	420
AACAAACCTG	CGTTGCTGCC	AGGGCGGCAA	ATGGGGAATG	GCGTCGTATC	AGCTACGCGG	480
AAATGTTCCA	CAACGTCCGC	GCCATCGCAC	AGAGCTTGCT	TCCTTACGGA	CTATCGGCAG	540
AGCGTCCGCT	GCTTATCGTC	TCTGGAAATG	ACCTGGAACA	TCTTCAGCTG	GCATT TGGGG	600
CTATGTATGC	GGGCATTCCC	TATTGCCCGG	TGTCTCCTGC	TTATTCACTG	CTGTCGCAAG	660
ATTTGGCGAA	GCTGCGTCAC	ATCGTAGGTC	TTCTGCAACC	GGGACTGGTC	TTTGTGCCG	720
ATGCGCAGC	TTTCCAGGGG	GAGAGGCGGT	TTGCGTATTC	CGGGCATGCA	TAAAAACTGT	780
TGTAATTCAT	TAAGCATTCT	GCCGACATGG	AAGCCATCAC	AAACGGCATG	ATGAACCTGA	840
ATCGCCAGCG	GCATCAGCAC	CTTGTCGCCT	TGCGTATAAT	ATTTGCCCAT	GGACGCACAC	900
CGTGGAACG	GATGAAGGCA	CGAACCCAGT	TGACATAAGC	CTGTTCCGGT	CGTAAACTGT	960
AATGCAAGTA	GCGTATGCGC	TCACGCAACT	GGTCCAGAAC	CTTGACCGAA	CGCAGCGGTG	1020
GTAACGGCGC	AGTGGCGGTT	TTCATGGCTT	GTTATGACTG	TTTTTTTGTG	CAGTCTATGC	1080
CTCGGGCATC	CAAGCAGCAA	GCGCGTTACG	CCGTGGGTCC	ATGTTTGATG	TTATGGAGCA	1140
GCAACGATGT	TACGCAGCAG	CAACGATGTT	ACGCAGCAGG	GCAGTCGCCC	TAAACAAAAG	1200
TTAGGTGGCT	CAAGTATGGG	CATCATTCGC	ACATGTAGGC	TCGGCCCTGA	CCAAGTCAAA	1260
TCCATGCGGG	CTGCTCTTGA	TCTTTTCCGGT	CGTGAGTTCG	GAGACGTAGC	CACCTACTCC	1320
CAACATCAGC	CGGACTCCGA	TTACCTCGGG	AACTTGCTCC	GTAAGTAAAC	ATTCATCGCG	1380
CTTGCTGCCT	TCGACCAAGA	AGCGGTTGTT	GGCGCTCTCG	CGGCTTACGT	TCTGCCCAGG	1440
TTTGAGCAGC	CGCGTAGTGA	GATCTATATC	TATGATCTCG	CAGTCTCCGG	CGAGCACCGG	1500
AGGCAGGGCA	TTGCCACCGC	GCTCATCAAT	CTCCTCAAGC	ATGAGGCCAA	CGCGCTTGGT	1560
GCTTATGTGA	TCTACGTGCA	AGCAGATTAC	GGTGACGATC	CCGCAGTGCC	TCTCTATACA	1620
AAGTTGGGCA	TACGGGAAGA	AGTGATGCAC	TTTGATATCG	ACCCAAGTAC	CGCCACCTAA	1680
CAATTCGTTC	AAGCCGAGAT	CGGCTTCCCC	TGTTTTGCAA	TGGCGGTCCG	CGAAAGTTGA	1740
TGCGCTGTAT	CGTGGTGAAG	ATCAATCCAT	GCTGCGTGAC	GAGGCCACAC	TGTGAGTTGG	1800
TCAGGGGGGG	CTTACTCGGC	GTTTTCCGAC	ACTGCGTTGG	TTGCGGAGT	GCGACCCCC	1860
TGGATTGATT	GCGGGGGTGC	CCTGTCGCTG	GTGTCGCCTA	TCGACTTAGG	GGTAAAGGTC	1920
GCTCGCGAAG	TTCTGATGCG	TGCGTCGCTT	GAACCACAAA	TGGTCGATAG	CGTACTCGCA	1980
GGCTCTATGG	CTCAAGCAAG	CTTTGATGCT	TACCTGCTCC	CGCGGCACAT	TGGCTTGTAC	2040
AGCGGTGTTT	CCAAGTCGGT	TCCGGCCTTG	GGGGTGCAGC	GCATTTGCGG	CACAGGCTTC	2100
GAAGTCTTTC	GGCAGGCCGG	CGAGCAGATT	TCCCAAGGCG	CTGATCACGT	GCTGTGTGTC	2160
GCGGGCTGCA	G					2171

Sequenz 9

CTGCAGCCGA	GCATCGATTG	AGCACTTTAC	CCAGCTGCGC	TGGCTGACCA	TTCAGAATGG	60
CCCCGCGCAC	TATCCAATCT	AAATCGATCT	TCGGGCGCCG	CGGGCATCAT	GCCCCGCGCG	120
CTCGCCTCAT	TTCAATCTCT	AACTTGATAA	AAACAGAGCT	GTTCTCCGGT	CTTGGTGGAT	180
CAAGGCCAGT	CGCGGAGAGT	CTCGAAGAGG	AGAGTACAGT	GAACGCCGAG	TCCACATTGC	240
AACCGCAGGC	ATCATCATGC	TCTGCTCAGC	CACGCTACCG	CAGTGTGTCG	ATTGGTCATC	300
CTCCGGTTGA	GGTTACGCAA	GACGCTGGAG	GTATTGTCCG	GATGCGTTCT	CTCGAGGCGC	360
TTCTTCCCTT	CCCCGGTCTGA	ATTCTTGAGC	GTCTCGAGCA	TTGGGCTAAG	ACCCGTCCAG	420
AACAAACCTG	CGTTGCTGCC	AGGGCGGCAA	ATGGGGAATG	GCGTCGTATC	AGCTACGCGG	480
AAATGTTCCA	CAACGTCCGC	GCCATCGCAC	AGAGCTTGCT	TCCTTACGGA	CTATCGGCAG	540
AGCGTCCGCT	GCTTATCGTC	TCTGGAAATG	ACCTGGAACA	TCTTCAGCTG	GCATTTGGGG	600
CTATGTATGC	GGGCATTCCC	TATTGCCCGG	TGTCTCCTGC	TTATTCACTG	CTGTCGCAAG	660
ATTTGGCGAA	GCTGCGTCAC	ATCGTAGGTC	TTCTGCAACC	GGGACTGGTC	TTTGCTGCCG	720
ATGCAGCACC	TTTCCAGCGC	GCTGTTTTGC	AATGGCGGTC	GGCGAAAGTT	GATGCGCTGT	780
ATCGTGGTGA	AGATCAATCC	ATGCTGCGTG	ACGAGGCCAC	ACTGTGAGTT	GGTCAGGGGG	840
GGCTTACTCG	GCGTTTTCCG	ACACTGCGTT	GGTTGCGGCA	GTGCGCACCC	CCTGGATTGA	900
TTGCGGGGGT	GCCCTGTCGC	TGGTGTGCGC	TATCGACTTA	GGGGTAAAGG	TCGCTCGCGA	960
AGTTCTGATG	CGTGCGTCGC	TTGAACCACA	AATGGTCGAT	AGCGTACTCG	CAGGCTCTAT	1020
GGCTCAAGCA	AGCTTTGATG	CTTACCTGCT	CCCGCGGCAC	ATTGGCTTGT	ACAGCGGTGT	1080
TCCCAAGTCG	GTTCCGGCCT	TGGGGGTGCA	GCGCATTTGC	GGCACAGGCT	TCGAACTGCT	1140
TCGGCAGGCC	GGCGAGCAGA	TTTCCCAAGG	CGCTGATCAC	GTGCTGTGTG	TCGCGGGCTG	1200
CAG						1203

Sequenz 10

GAATTCCTT	GGCGACGAAA	GGGCGGCAGG	CCGCATGGCC	ACGGCTGGGC	GGTAACTGAT	60
GCTTGCGTTA	ATCGTTAACC	GTTTGAAATT	CCTTGCCAAA	TTTCGGCGAG	AGAATCATGC	120
GGGTACGCTT	TTCCGTGCGC	TTTGATCTGC	GCTTCCGTGC	CTTGAATCAG	AAAAATAGTT	180
AATTGACAGA	ACTATAGGTT	CGCAGTAGCT	TTTGCTCACC	CACCAAATCC	ACAGCACTGG	240
GGTGACGAT	GAATAGCTAC	GATGGCCGTT	GGTCTACCGT	TGATGTGAAG	GTTGAAGAAG	300
GTATCGCTTG	GGTCACGCTG	AACCGCCCGG	AGAAGCGCAA	CGCAATGAGC	CCAACTCTCA	360
ATCGAGAGAT	GGTCGAGGTT	CTGGAGGTGC	TGGAGCAGGA	CGCAGATGCT	CGCGTGCTTG	420
TTCTGACTGG	TGCAGGCGAA	TCCTGGACCG	CGGGCATGGA	CCTGAAGGAG	TATTTCCGCG	480
AGACCGATGC	TGGCCCCGAA	ATTCTGCAAG	AGAAGATTCG	TCGGGGACAG	CAAGCGAACC	540
GGAATTGCCA	GCTGGGGGCG	CCTCTGGTAA	GGTTGGGAAG	CCCTGCAAAG	TAAACTGGAT	600
GGCTTTCTTG	CCGCCAAGGA	TCTGATGGCG	CAGGGGATCA	AGATCTGATC	AAGAGACAGG	660
ATGAGGATCG	TTTCGCATGA	TTGAACAAGA	TGGATTGCAC	GCAGGTTCTC	CGGCCGCTTG	720
CGTGGAGAGG	CTATTCCGGT	ATGACTGCGC	ACAAGAGACA	ATCGGCTGCT	CTGATGCCCC	780
CGTGTTCCGG	CTGTACGCGC	AGGGGCGCCC	GGTTCTTTTT	GTCAAGACCG	ACCTGTCCGG	840
TGCCCAGAAT	GAAGTGCAGG	ACGAGGCAGC	GCGGCTATCG	TGGCTGGCCA	CGACGGGCGT	900
TCCTTGCGCA	GCTGTGCTCG	ACGTTGTAC	TGAAGCGGGA	AGGGACTGGC	TGCTATTGGG	960
CGAAGTGCCG	GGGCAGGATC	TCCTGTCTATC	TCACCTTGCT	CCTGCCGAGA	AAGTATCCAT	1020
CATGGCTGAT	GCAATGCGGC	GGCTGCATAC	GCTTGATCCG	GCTACCTGCC	CATTGACCA	1080
CCAAGCGAAA	CATCGCATCG	AGCGAGCACG	TACTCGGATG	GAAGCCGGTC	TTGTGATCA	1140
GGATGATCTG	GACGAAGAGC	ATCAGGGGCT	CGCGCCAGCC	GAAGTGTTCG	CCAGGCTCAA	1200
GGCGCGCATG	CCCGACGGCG	AGGATCTCGT	CGTGACCCAT	GGCGATGCCT	GCTTGCCGAA	1260
TATCATGGTG	GAAAATGGCC	GCTTTTCTGG	ATTCATCGAC	TGTGGCCGGC	TGGGTGTGGC	1320
GGACCGCTAT	CAGGACATAG	CGTTGGCTAC	CCGTGATATT	GCTGAAGAGC	TTGGCGGCGA	1380
ATGGGCTGAC	CGCTTCTCG	TGCTTTACGG	TATCGCCGCT	CCCGATTTCG	AGCGCATCGC	1440
CTTCTATCGC	CTTCTTGACG	AGTTCTTCTG	AGCGGGACTC	TGGGGTTTCA	AATGACCGAC	1500
CAAGCGACGC	CCCGAGCAGG	GCATGAAGCA	GTTCTTTGAC	GAGAAAAGCA	TCAAGCCGGG	1560
CTTGACAGAC	TACAAGCGCT	GATAAATGCG	CCGGGGCCCT	CGCTGCGCCC	CCGGCCTTCC	1620
AATAATGACA	ATAATGAGGA	GTGCCCCAATG	TTTCACGTGC	CCCTGCTTAT	TGGTGGTAAG	1680
CCTTGTTTCA	CATCTGATGA	GCGCACCTTC	GAGCGTCGTA	GCCCCTGAC	CGGAGAAGTG	1740
GTATCGCGCG	TCGCTGCTGC	CAGTTTGGA	GATGCGGACG	CCGCAGTGGC	CGCTGCACAG	1800
GCTGCGTTTC	CTGAATGGGC	GGCGCTTGCT	CCGAGCGAAC	GCCGTGCCCC	ACTGCTGCGA	1860
GCGGCGGATC	TTCTAGAGGA	CCGTCTTCC	GAGTTCACCG	CCGCAGCGAG	TGAAACTGGC	1920
GCAGCGGGAA	ACTGGTATGG	GTTTAACGTT	TACCTGGCGG	CGGGCATGTT	GCGGGGAATT	1980
C						1981

Sequenz 11

GAATTC	CCCT	GGCGACGAAA	GGGCGGCAGG	CCGCATGGCC	ACGGCTGGGC	GGTAACTGAT	60
GCTTGC	GTTA	ATCGTTAACC	GTTTGAAATT	CCTTGCCAAA	TTTCGGCGAG	AGAATCATGC	120
GGGTAC	GCCT	TTCCGTGCGC	TTTGATCTGC	GCTTCCGTGC	CTTGAATCAG	AAAAATAGTT	180
AATTGAC	AGATA	CTATAGGTT	CGCAGTAGCT	TTTGCTCACC	CACCAAATCC	ACAGCACTGG	240
GGTGCAC	GAT	GAATAGCTAC	GATGGCCGTT	GGTCTACCGT	TGATGTGAAG	GTTGAAGAAG	300
GTATCG	CTTG	GGTCACGCTG	AACCGCCCCG	AGAAGCGCAA	CGCAATGAGC	CCAACCTCTCA	360
ATCGAG	AGAT	GGTCGAGGTT	CTGGAGGTGC	TGGAGCAGGA	CGCAGATGCT	CGCGTGCTTG	420
TTCTGAC	TGG	TGCAGGCGAA	TCCTGGACCG	CGGGCATGGA	CCTGAAGGAG	TATTTCCGCG	480
AGACCG	ATGC	TGGCCCCGAA	ATTCTGCAAG	AGAAGATTCTG	TCGGGGGAGA	GGCGGTTTGC	540
GTATTGG	GCG	CATGCATAAA	AACTGTTGTA	ATTCAATTAAG	CATTCTGCCG	ACATGGAAGC	600
CATCACA	AAAC	GGCATGATGA	ACCTGAATCG	CCAGCGGCAT	CAGCACCTTG	TCGCCTTGCG	660
TATAAT	ATTT	GCCCATGGAC	GCACACCGTG	GAAACGGATG	AAGGCACGAA	CCCAGTTGAC	720
ATAAGCG	TGT	TGGCTTCGTA	AACTGTAATG	CAAGTAGCGT	ATGCGCTCAC	GCAACTGGTC	780
CAGAAC	CTTG	ACCGAACGCA	GCGGTGGTAA	CGGCGCAGTG	GCGGTTTTCA	TGGCTTGTTA	840
TGACTG	TTTT	TTTGTACAGT	CTATGCCTCG	GGCATCCAAG	CAGCAAGCGC	GTTACGCCGT	900
GGGT	CGATGT	TTGATGTTAT	GGAGCAGCAA	CGATGTTACG	CAGCAGCAAC	GATGTTACGC	960
AGCAGG	GCGAG	TCGCCCTAAA	ACAAAGTTAG	GTGGCTCAAG	TATGGGCATC	ATTTCGCACAT	1020
GTAGG	CTCGG	CCCTGACCAA	GTCAAATCCA	TGCGGGCTGC	TCTTGATCTT	TTCGGTCGTG	1080
AGTTCG	GAGA	CGTAGCCACC	TACTCCCAAC	ATCAGCCGGA	CTCCGATTAC	CTCGGGAAC	1140
TGCTCC	GTAG	TAAGACATTC	ATCGCGCTTG	CTGCCTTCGA	CCAAGAAGCG	GTTGTTGGCG	1200
CTCTCG	GCGC	TTACGTTCTG	CCCAGGTTTG	AGCAGCCGCG	TAGTGAGATC	TATATCTATG	1260
ATCTCG	CAGT	CTCCGGCGAG	CACCGGAGGC	AGGGCATTCG	CACCGCGCTC	ATCAATCTCC	1320
TCAAG	CATGA	GGCCAACGCG	CTTGGTGCTT	ATGTGATCTA	CGTGCAAGCA	GATTACGGTG	1380
ACGATC	CCCG	AGTGGCTCTC	TATACAAAGT	TGGGCATACG	GGAAGAAGTG	ATGCACTTTG	1440
ATATCG	ACCC	AAGTACCGCC	ACCTAACAAT	TCGTTCAAGC	CGAGATCGGC	TTCCCCGAGC	1500
AGGGC	CATGAA	GCAGTTCCCTT	GACGAGAAAA	GCATCAAGCC	GGGCTTGCGAG	ACCTACAAGC	1560
GCTGATA	AAAT	GCGCCGGGGC	CCTCGCTGCG	CCCCGGCCCT	TCCAATAATG	ACAATAATGA	1620
GGAGT	GCCCCA	ATGTTTCACG	TGCCCCTGCT	TATTGGTGGT	AAGCCTTGTT	CAGCATCTGA	1680
TGAGCG	CACC	TTCGAGCGTC	GTAGCCCGCT	GACCGGAGAA	GTGGTATCGC	GCGTCGCTGC	1740
TGCCAG	TTTG	GAAGATGCGG	ACGCCGAGT	GGCCGCTGCA	CAGGCTGCGT	TTCTTGAATG	1800
GGCGG	CGCTT	GCTCCGAGCG	AACGCCGTGC	CCGACTGCTG	CGAGCGGCGG	ATCTTCTAGA	1860
GGACCG	TTCT	TCCGAGTTCA	CCGCCGAGC	GAGTGAAACT	GGCGCAGCGG	GAAACTGGTA	1920
TGGGTT	TTAAC	GTTTACCTGG	CGGCGGGCAT	GTTGCGGGGA	ATTC		1964

Sequenz 12

GAATTCCCCT	GGCGACGAAA	GGGCGGCAGG	CCGCATGGCC	ACGGCTGGGC	GGTAACTGAT	60
GCTTGCCTTA	ATCGTTAACC	GTTTGAAATT	CCTTGCCAAA	TTTCGGCGAG	AGAATCATGC	120
GGGTACGCCT	TTCCGTGCGC	TTTGATCTGC	GCTTCCGTGC	CTTGAATCAG	AAAAATAGTT	180
AATTGACAGA	ACTATAGGTT	CGCAGTAGCT	TTTGCTCACC	CACCAAATCC	ACAGCACTGG	240
GGTGACGAT	GAATAGCTAC	GATGGCCGTT	GGTCTACCGT	TGATGTGAAG	GTTGAAGAAG	300
GTATCGCTTG	GGTCACGCTG	AACCGCCCGG	AGAAGCGCAA	CGCAATGAGC	CCAACTCTCA	360
ATCGAGAGAT	GGTCGAGGTT	CTGGAGGTGC	TGGAGCAGGA	CGCAGATGCT	CGCGTGCTTG	420
TTCTGACTGG	TGCAGGCGAA	TCCTGGACCG	CGGGCATGGA	CCTGAAGGAG	TATTTCCGCG	480
AGACCGATGC	TGGCCCCGAA	ATTCTGCAAG	AGAAGATTCTG	TCGCGAGCAG	GGCATGAAGC	540
AGTTCCTTGA	CGAGAAAAGC	ATCAAGCCGG	GCTTGCAGAC	CTACAAGCGC	TGATAAATGC	600
GCCGGGGCCC	TCGCTGCGCC	CCCGGCCTTC	CAATAATGAC	AATAATGAGG	AGTGCCCAAT	660
GTTTCACGTG	CCCCTGCTTA	TTGGTGGTAA	GCCTTGTTCA	GCATCTGATG	AGCGCACCTT	720
CGAGCGTCGT	AGCCCGCTGA	CCGGAGAAGT	GGTATCGCGC	GTCGCTGCTG	CCAGTTTGGA	780
AGATGCGGAC	GCCGCAGTGG	CCGCTGCACA	GGCTGCGTTT	CCTGAATGGG	CGGCGCTTGC	840
TCCGAGCGAA	CGCCGTGCCC	GACTGCTGCG	AGCGGCGGAT	CTTCTAGAGG	ACCGTTCTTC	900
CGAGTTCACC	GCCGCAGCGA	GTGAAACTGG	CGCAGCGGGA	AACTGGTATG	GGTTTAACGT	960
TTACCTGGCG	GCGGGCATGT	TGCGGGGAAT	TC			992

Sequenz 13

GAATTCCAAT	AATGACAATA	ATGAGGAGTG	CCCAATGTTT	CACGTGCCCC	TGCTTATTGG	60
TGGTAAGCCT	TGTTTACGAT	CTGATGAGCG	CACCTTCGAG	CGTCGTAGCC	CGCTGACCGG	120
AGAAAGTGTA	TCGCGCGTCG	CTGCTGCCAG	TTTGGAAGAT	GCGGACGCCG	CAGTGGCCCG	180
TGCACAGGCT	GCGTTTCCTG	AATGGGCGGC	GCTTGCTCCG	AGCGAACGCC	GTGCCCAGCT	240
GCTGCGAGCG	GCGGATCTTC	TAGAGGACCG	TTCTTCCGAG	TTCACCGCCG	CAGCGAGTGA	300
AACTGGCGCA	GCGGGAAACT	GGTATGGGTT	TAACGTTTAC	CTGGCGGCCG	GCATGTTGCG	360
GGAAGCCGCG	GCCATGACCA	CACAGATTCA	GGGCGATGTC	ATTCCGTCCA	ATGTGCCCGG	420
TAGCTTTGCC	ATGGCGGTTT	GACAGCCATG	TGGCGTGCTG	CTCGGTATTG	CGCCTTGGA	480
TGCTCCGGTA	ATCCTTGGCG	TACGGGCTGT	TGCGATGCCG	TTGGCATGCG	GCAATACCGT	540
GGTGTTGAAA	AGCTCTGAGC	TGAGTCCCTT	TACCCATCGC	CTGATTGGTC	AGGTGTTGCA	600
TGATGCTGGT	CTGGGGGATG	GCGTGGTGAA	TGTCATCAGC	AATGCCCCGC	AAGACGCTCC	660
TGCGGTGGTG	GAGCGACTGA	TTGCAAATCC	TGCGGTACGT	CGAGTGAAGT	TCACCGGTTT	720
GACCGACGTT	GGACGGATCA	TTGGTGAGCT	GTCTGCGCGT	CATCTGAAGC	CTGCTGTGCT	780
GGAATTAGGT	GGTAAGGCTC	CGTTCTTGGT	CTTGGACGAT	GCCGACCTCG	ATGCGGCGGT	840
CGAAGCGGCG	GCCTTTGGTG	CCTACTTCAA	TCAGGGTCAA	ATCTGCATGT	CCACTGAGCG	900
TCTGATTGTG	ACAGCAGTCG	CAGACGCCTT	TGTTGAAAAG	CTGGCGAGGA	AGGTCGCCAC	960
ACTGCGTGCT	GGCGATCCTA	ATGATCCGCA	ATCGGTCTTG	GGTTCGTTGA	TTGATGCCAA	1020
TGCAGGTCAA	CGCATCCAGG	TTCTGGTCTG	TGATGCGCTC	GGGGACAGCA	AGCGAACCGG	1080
AATTGCCAGC	TGGGGCGCCC	TCTGGTAAGG	TTGGGAAGCC	CTGCAAAGTA	AACTGGATGG	1140
CTTTCTTGCC	GCCAAGGATC	TGATGGCGCA	GGGGATCAAG	ATCTGATCAA	GAGACAGGAT	1200
GAGGATCGTT	TCGCATGATT	GAACAAGATG	GATTGCACGC	AGGTTCTCCG	GCCGCTTGGG	1260
TGGAGAGGCT	ATTCGGCTAT	GACTGGGCAC	AACAGACAAT	CGGCTGCTCT	GATGCCGCCG	1320
TGTTCCGGCT	GTCAGCGCAG	GGGCGCCCGG	TTCTTTTGTG	CAAGACCGAC	CTGTCCGGTG	1380
CCCTGAATGA	ACTGCAGGAC	GAGGCAGCGC	GGCTATCGTG	GCTGGCCACG	ACGGGCGTTC	1440
CTTGCGCAGC	TGTGCTCGAC	GTTGTCACTG	AAGCGGGAAG	GGACTGGCTG	CTATTGGGCG	1500
AAGTGCCGGG	GCAGGATCTC	CTGTCTATCT	ACCTTGCTCC	TGCCGAGAAA	GTATCCATCA	1560
TGGCTGATGC	AATGCGGCGG	CTGCATACGC	TTGATCCGGC	TACCTGCCCA	TTCGACCACC	1620
AAGCGAAACA	TCGCATCGAG	CGAGCACGTA	CTCGGATGGA	AGCCGGTCTT	GTCGATCAGG	1680
ATGATCTGGA	CGAAGAGCAT	CAGGGGCTCG	CGCCAGCCGA	ACTGTTCCGC	AGGCTCAAGG	1740
CGCGCATGCC	CGACGGCGAG	GATCTCGTCG	TGACCCATGG	CGATGCCTGC	TTGCCGAATA	1800
TCATGGTGGA	AAATGGCCGC	TTTTCTGGAT	TCATCGACTG	TGGCCGGCTG	GGTGTGGCGG	1860
ACCGCTATCA	GGACATAGCG	TTGGCTACCC	GTGATATTGC	TGAAGAGCTT	GGCGGCGAAT	1920
GGGCTGACCG	CTTCTCGTG	CTTTACGGTA	TCGCCGCTCC	CGATTTCGAG	CGCATCGCCT	1980
TCTATCGCCT	TCTTGACGAG	TTCTTCTGAG	CGGGACTCTG	GGGTTTCGAA	TGACCGACCA	2040
AGCGACGCCC	GGCCCAGCGC	GTCGATTTCG	GCATTTGCCA	TATCAATGGA	CCGACTGTGC	2100
ATGACGAGGC	TCAGATGCCA	TTCGGTGGGG	TGAAGTCCAG	CGGCTACGGC	AGCTTCGGCA	2160

GTCGAGCATC	GATTGAGCAC	TTTACCCAGC	TGCGCTGGCT	GACCATTGAG	AATGGCCCGC	2220
GGCACTATCC	AATCTAAATC	GATCTTCGGG	CGCCGCGGGC	ATCATGCCCC	CGGCGCTCGC	2280
CTCATTTCAA	TCTCTAACTT	GATAAAAACA	GAGCTGTTCT	CCGGTCTTGG	TGGATCAAGG	2340
CCAGTCGCGG	AGAGTCTCGA	AGAGGAGAGT	ACAGTGAACG	CCGAGTCCAC	ATTGCAACCG	2400
CAGGCATCAT	CATGCTCTGC	TCAGCCACGC	TACCGCAGTG	TGTCGATTGG	TCATCCTCCG	2460
GTTGAGGTTA	CGCAAGACGC	TGGAGGTATT	GTCCGGATGC	GTTCTCTCGA	GGCGCTTCTT	2520
CCCTTCCCGG	GTGGAATTC					2539

Sequenz 14

GAATTCCAAT	AATGACAATA	ATGAGGAGTG	CCCAATGTTT	CACGTGCCCC	TGCTTATTGG	60
TGGTAAGCCT	TGTTTCAGCAT	CTGATGAGCG	CACCTTCGAG	CGTCGTAGCC	CGCTGACCGG	120
AGAAGTGGTA	TCGCGCGTCG	CTGCTGCCAG	TTTGAAGAT	GCGGACGCCG	CAGTGGCCGC	180
TGCACAGGCT	GCGTTTCCTG	AATGGGCGGC	GCTTGCTCCG	AGCGAACGCC	GTGCCCGACT	240
GCTGCGAGCG	GCGGATCTTC	TAGAGGACCG	TTCTTCCGAG	TTCACCGCCG	CAGCGAGTGA	300
AACTGGCGCA	GCGGAAACT	GGTATGGGTT	TAACGTTTAC	CTGGCGGCGG	GCATGTTGCG	360
GGAAAGCCGC	GCCATGACCA	CACAGATTCA	GGGCGATGTC	ATTCCGTCCA	ATGTGCCCGG	420
TAGCTTTGCC	ATGGCGGTTT	GACAGCCATG	TGGCGTGTTG	CTCGGTATTG	CGCCTTGGAA	480
TGCTCCGGTA	ATCCTTGGCG	TACGGGCTGT	TGCGATGCCG	TTGGCATGCG	GCAATACCGT	540
GGTGTGAAA	AGCTCTGAGC	TGAGTCCCTT	TACCCATCGC	CTGATTGGTC	AGGTGTTGCA	600
TGATGCTGGT	CTGGGGGATG	GCGTGGTGAA	TGTCATCAGC	AATGCCCCGC	AAGACGCTCC	660
TGCGGTGGTG	GAGCGACTGA	TTGCAAATCC	TGCGGTACGT	CGAGTGAAGT	TCACCGGTTT	720
GACCGAGGTT	GGACGGATCA	TTGGTGAGCT	GTCTGCGCGT	CATCTGAAGC	CTGCTGTGCT	780
GGAAATTAGGT	GGTAAGGCTC	CGTTCTTGGT	CTTGACGAT	GCCGACCTCG	ATGCGGCGGT	840
CGAAGCGGCG	GCCTTTGGTG	CCTACTTCAA	TCAGGGTCAA	ATCTGCATGT	CCACTGAGCG	900
TCTGATTGTG	ACAGCAGTCG	CAGACGCCCT	TGTTGAAAAG	CTGGCGAGGA	AGGTCGCCAC	960
ACTGCGTGCT	GGCGATCCTA	ATGATCCGCA	ATCGGTCTTG	GGTTCGTTGA	TTGATGCCAA	1020
TGCAGGTCAA	CGCATCCAGG	TGGGGAGAGG	CGGTTTGCGT	ATTGGGCGCA	TGCATAAAAA	1080
CTGTTGTAAT	TCATTAAGCA	TTCTGCCGAC	ATGGAAGCCA	TCACAAACGG	CATGATGAAC	1140
CTGAATCGCC	AGCGGCATCA	GCACCTTGTC	GCCTTGCGTA	TAATATTTGC	CCATGGACGC	1200
ACACCGTGGA	AACGGATGAA	GGCACGAACC	CAGTTGACAT	AAGCCTGTTC	GGTTCGTAAA	1260
CTGTAATGCA	AGTAGCGTAT	GCGCTCACGC	AACTGGTCCA	GAACCTTGAC	CGAACGCAGC	1320
GGTGTAACG	GCGCAGTGCC	GGTTTTTCATG	GCTTGTTATG	ACTGTTTTTTT	TGTACAGTCT	1380
ATGCCTCGGG	CATCCAAGCA	GCAAGCGCGT	TACGCCGTGG	GTCGATGTTT	GATGTTATGG	1440
AGCAGCAACG	ATGTTACGCA	GCAGCAACGA	TGTTACGCAG	CAGGGCAGTC	GCCCTAAAAC	1500
AAAGTTAGGT	GGCTCAAGTA	TGGGCATCAT	TCGCACATGT	AGGCTCGGCC	CTGACCAAGT	1560
CAAATCCATG	CGGGCTGCTC	TTGATCTTTT	CGGTCGTGAG	TTCGGAGACG	TAGCCACCTA	1620
CTCCCAACAT	CAGCCGGACT	CCGATTACCT	CGGGAACTTG	CTCCGTAGTA	AGACATTTCAT	1680
CGCGCTTGCT	GCCTTCGACC	AAGAAGCGGT	TGTTGGCGCT	CTCGCGGCTT	ACGTTCTGCC	1740
CAGGTTTGAG	CAGCCGCGTA	GTGAGATCTA	TATCTATGAT	CTCGCAGTCT	CCGGCGAGCA	1800
CCGGAGGCAG	GGCATTGCCA	CCGCGCTCAT	CAATCTCCTC	AAGCATGAGG	CCAACGCGCT	1860
TGGTGCTTAT	GTGATCTACG	TGCAAGCAGA	TTACGGTGAC	GATCCCGCAG	TGGCTCTCTA	1920
TACAAAGTTG	GGCATACGGG	AAGAAGTGAT	GCACTTTGAT	ATCGACCCAA	GTACCGCCAC	1980
CTAACAAATC	GTTCAAGCCG	AGATCGGCTT	CCCAATTGGC	CCAGCGCGTC	GATTCGGGCA	2040
TTTGCCATAT	CAATGGACCG	ACTGTGCATG	ACGAGGCTCA	GATGCCATTTC	GGTGGGGTGA	2100
AGTCCAGCGG	CTACGGCAGC	TTCGGCAGTC	GAGCATCGAT	TGAGCACTTT	ACCCAGCTGC	2160

GCTGGCTGAC	CATTCAGAAT	GGCCCGCGGC	ACTATCCAAT	CTAAATCGAT	CTTCGGGCGC	2220
CGCGGGCATC	ATGCCC GCGG	CGCTCGCCTC	ATTTCAATCT	CTAACTTGAT	AAAAACAGAG	2280
CTGTTCTCCG	GTCTTG GTGG	ATCAAGGCCA	GTCGCGGAGA	GTCTCGAAGA	GGAGAGTACA	2340
GTGAACGCCG	AGTCCACATT	GCAACCGCAG	GCATCATCAT	GCTCTGCTCA	GCCACGCTAC	2400
CGCAGTGTGT	CGATTGGTCA	TCCTCCGGTT	GAGGTTACGC	AAGACGCTGG	AGGTATTGTC	2460
CGGATGCGTT	CTCTCGAGGC	GCTTCTTCCC	TTCCCGGGTG	GAATTC		2506

Sequenz 15

GAATTCCAAT	AATGACAATA	ATGAGGAGTG	CCCAATGTTT	CACGTGCCCC	TGCTTATTGG	60
TGGTAAGCCT	TGTTCAGCAT	CTGATGAGCG	CACCTTCGAG	CGTCGTAGCC	CGCTGACCGG	120
AGAAGTGGTA	TCGCGCGTCG	CTGCTGCCAG	TTTGGAAGAT	GCGGACGCCG	CAGTGGCCCG	180
TGCACAGGCT	GCGTTTCCTG	AATGGGCGGC	GCTTGCTCCG	AGCGAACGCC	GTGCCCAGCT	240
GCTGCGAGCG	GCGGATCTTC	TAGAGGACCG	TTCTTCCGAG	TTCAACCGCCG	CAGCGAGTGA	300
AACTGGCGCA	GCGGGAAACT	GGTATGGGTT	TAACGTTTAC	CTGGCGGCGG	GCATGTTGCG	360
GGAAGCCGCG	GCCATGACCA	CACAGATTCA	GGGCGATGTC	ATTCCGTCCA	ATGTGCCCGG	420
TAGCTTTGCC	ATGGCGGTTC	GACAGCCATG	TGGCGTGGTG	CTCGGTATTG	CGCCTTGGAA	480
TGCTCCGGTA	ATCCTTGGCG	TACGGGCTGT	TGCGATGCCG	TTGGCATGCG	GCAATACCGT	540
GGTGTTGAAA	AGCTCTGAGC	TGAGTCCCTT	TACCCATCGC	CTGATTGGTC	AGGTGTTGCA	600
TGATGCTGGT	CTGGGGGATG	GCGTGGTGAA	TGTCATCAGC	AATGCCCCGC	AAGACGCTCC	660
TGCGGTGGTG	GAGCGACTGA	TTGCAAATCC	TGCGGTACGT	CGAGTGAAC	TCACCGGTTT	720
CACCCAGCTT	GCAGCCATCA	TTGCTGAGCT	CTCTGCGCCT	CATCTGAAGC	CTGCTGTGCT	780
GGAATTAGGT	GGTAAGGCTC	CGTTCTTGGT	CTTGGACGAT	GCCGACCTCG	ATGCGGCGGT	840
CGAAGCGGCG	GCCTTTGGTG	CCTACTTCAA	TCAGGGTCAA	ATCTGCATGT	CCACTGAGCG	900
TCTGATTGTG	ACAGCAGTCG	CAGACGCCTT	TGTTGAAAAG	CTGGCGAGGA	AGGTGCGCCAC	960
ACTGCGTGCT	GGCGATCCTA	ATGATCCGCA	ATCGGTCTTG	GGTTCGTTGA	TTGATGCCAA	1020
TGCAGGTCAA	CGCATCCAGG	TTCTGGTCTGA	TGATGCGCTC	GCAAAAGGCG	CGCAATGGAA	1080
TTGGCCCAGC	GCGTCGATTC	GGGCATTTGC	CATATCAATG	GACCGACTGT	GCATGACGAG	1140
GCTCAGATGC	CATTCGGTGG	GGTGAAGTCC	AGCGGCTACG	GCAGCTTCGG	CAGTCGAGCA	1200
TCGATTGAGC	ACTTTACCCA	GCTGCGCTGG	CTGACCATTG	AGAATGGCCC	GCGGCACTAT	1260
CCAATCTAAA	TCGATCTTCG	GGCGCCGCGG	GCATCATGCC	CGCGGCGCTC	GCCTCATTTT	1320
AATCTCTAAC	TTGATAAAAA	CAGAGCTGTT	CTCCGGTCTT	GGTGGATCAA	GGCCAGTCGC	1380
GGAGAGTCTC	GAAGAGGAGA	GTACAGTGAA	CGCCGAGTCC	ACATTGCAAC	CGCAGGCATC	1440
ATCATGCTCT	GCTCAGCCAC	GCTACCGCAG	TGTGTCGATT	GGTCATCCTC	CGGTTGAGGT	1500
TACGCAAGAC	GCTGGAGGTA	TTGTCCGGAT	GCGTTCTCTC	GAGGCGCTTC	TTCCCTTCCC	1560
GGGTGGAATT	C					1571

Sequenz 16

GAATTCCGCG	GTCGGCGAAA	GTTGATGCGC	TGTATCGTGG	TGAAGATCAA	TCCATGCTGC	60
GTGACGAGGC	CACACTGTGA	GTTGGTCAGG	GGGGGCTTAC	TCGGCGTTTT	CCGACACTGC	120
GTTGGTTGCG	GCACTGCGCA	CCCCCTGGAT	TGATTGCGGG	GGTGCCCTGT	CGCTGGTGTG	180
GCCTATCGAC	TTAGGGGTAA	AGGTCGCTCG	CGAAGTTCTG	ATGCGTGCGT	CGCTTGAACC	240
ACAAATGGTC	GATAGCGTAC	TCGCAGGCTC	TATGGCTCAA	GCAAGCTTTG	ATGCTTACCT	300
GCTCCCGCGG	CACATTGGCT	TGTACAGCGG	TGTTCCCAAG	TCGGTTCCGG	CCTTGGGGGT	360
GCAGCGCATT	TGCGGCACAG	GCTTCGAACT	GCTTCGGCAG	GCCGGCGAGC	AGATTTCCTA	420
AGGCGCTGAT	CACGTGCTGT	GTGTCGCGGC	AGAGTCCATG	TCGCGTAACC	CCATCGCGTC	480
GTATACACAC	CGGGGCGGGT	TCCGCCTCGG	TGCGCCCGTT	GAGTTCAAGG	ATTTTTTGTG	540
GGAGGCATTG	TTTGATCCTG	CTCCAGGACT	CGACATGATC	GCTACCGCAG	AAAACCTGGG	600
GACAGCAAGC	GAACCGGAAT	TGCCAGCTGG	GGCGCCCTCT	GGTAAGGTTG	GGAAGCCCTG	660
CAAAGTAAAC	TGGATGGCTT	TCTTGCCGCC	AAGGATCTGA	TGGCGCAGGG	GATCAAGATC	720
TGATCAAGAG	ACAGGATGAG	GATCGTTTCG	CATGATTGAA	CAAGATGGAT	TGCACGCGAG	780
TTCTCCGGCC	GCTTGGGTGG	AGAGGCTATT	CGGCTATGAC	TGGGCACAAC	AGACAATCGG	840
CTGCTCTGAT	GCCGCCGTGT	TCCGGCTGTC	AGCGCAGGGG	CGCCCGGTTC	TTTTTGTCAA	900
GACCGACCTG	TCCGGTGCCC	TGAATGAACT	GCAGGACGAG	GCAGCGCGGC	TATCGTGGCT	960
GGCCACGACG	GGCGTTCTTT	GCGCAGCTGT	GCTCGACGTT	GTCAGTGAAG	CGGGAAGGGA	1020
CTGGCTGCTA	TTGGGCGAAG	TGCCGGGGCA	GGATCTCCTG	TCATCTCACC	TGCTCTCTGC	1080
CGAGAAAGTA	TCCATCATGG	CTGATGCAAT	GCGGCGGCTG	CATACGCTTG	ATCCGGCTAC	1140
CTGCCCATTG	GACCACCAAG	CGAAACATCG	CATCGAGCGA	GCACGTACTC	GGATGGAAGC	1200
CGGTCTTGTC	GATCAGGATG	ATCTGGACGA	AGAGCATCAG	GGGCTCGCGC	CAGCCGAACT	1260
GTTGCGCCAGG	CTCAAGGCGC	GCATGCCCGA	CGGCGAGGAT	CTCGTCGTGA	CCCATGGCGA	1320
TGCCTGCTTG	CCGAATATCA	TGGTGGAATA	TGGCCGCTTT	TCTGGATTCA	TCGACTGTGG	1380
CCGGCTGGGT	GTGGCGGACC	GCTATCAGGA	CATAGCGTTG	GCTACCCGTG	ATATTGCTGA	1440
AGAGCTTGCG	GGCGAATGGG	CTGACCGCTT	CCTCGTGCTT	TACGGTATCG	CCGCTCCCGA	1500
TTGCGAGCGC	ATCGCCTTCT	ATCGCCTTCT	TGACGAGTTC	TTCTGAGCGG	GACTCTGGGG	1560
TTGCAAAATGA	CCGACCAAGC	GACGCCCATG	GAGGGCGCAA	GAGGAGAAAT	GGATTGACCA	1620
AGAGATCGTG	GCTGTTACGG	ATGAACAGTT	CGATTTAGAG	GGCTACAACA	GTCGAGCAAT	1680
TGAACTGCCT	CGGAAGGCAA	AATTGTTGAT	CGTGACAGTC	ATCCGCGGCC	TAGCAGTCTT	1740
TGAAGCCCTT	TCCCATTGTA	AGCCTGTTCA	TTCTGGCGGG	GTGCAGACTG	CGGGCAACAG	1800
CTGTGCCGTA	GTGGACGGCG	CCGCGGCGGC	TTTGGTGGCT	CGAGAGTCGT	CTGCGACACA	1860
GCCGGTCTTG	GCTAGGATAC	TGGCTACCTC	CGTAGTCGGG	ATCGAGCCCG	AGCATATGGG	1920
GCTCGGCCCT	GCGCCGCGA	TTGCGCTGCT	GCTTGCGCGT	AGTGATCTTA	GTTTGAGGGA	1980
TATCGACCTC	TTTGAGATAA	ACGAGGCGCA	GGCCGCCCAA	GTTCTAGCGG	TACAGCATGA	2040
ATTGGGTATT	GAGCACTCAA	AACTTAATAT	TTGGGGCGGG	GCCATTGCAC	TTGGACACCC	2100
GCTTGCCGCG	ACCGGATTGC	GTCTCTGCAT	GACCCTCGCT	CACCAATTGC	AAGCTAATAA	2160

CTTTCGATAT	GGAATTGCCT	CGGCATGCAT	TGGTGGGGGA	CAGGGGATGG	CGGTTCTTTT	2220
AGAGAATCCC	CACTTCGGTT	CGTCCTCTGC	ACGAAGTTCG	ATGATTAACA	GAGTTGACCA	2280
CTATCCACTG	AGCTAACGGG	CATCTCCTTT	GTTGCTTTGA	GGTGGCGCAC	GAAGGAGGGC	2340
TCGAAAATCT	CTGCTAAAAA	CAAGAAGAAG	GAACAGGGAA	CATGATTAGT	TTCGCTCGTA	2400
TGGCAGAAAG	TTTAGGAGTC	CAGGCTAAAC	TTGCCCTTGC	CTTCGCACTC	GTATTATGTG	2460
TCGGGCTGAT	TGTTACCGGC	ACGGGTTTCT	ACAGTGTACA	TACCTTGTC	GGGTTGGTGG	2520
GAATTC						2526

Sequenz 17

GAATTCCGCG	GTCGGCGAAA	GTTGATGCGC	TGTATCGTGG	TGAAGATCAA	TCCATGCTGC	60
GTGACGAGGC	CACACTGTGA	GTTGGTCAGG	GGGGGCTTAC	TCGGCGTTTT	CCGACACTGC	120
GTTGGTTGCG	GCAGTGCGCA	CCCCCTGGAT	TGATTGCGGG	GGTGCCCTGT	CGCTGGTGTC	180
GCCTATCGAC	TTAGGGGTAA	AGGTCGCTCG	CGAAGTTCTG	ATGCGTGCGT	CGCTTGAACC	240
ACAAATGGTC	GATAGCGTAC	TCGCAGGCTC	TATGGCTCAA	GCAAGCTTTG	ATGCTTACCT	300
GCTCCCGCGG	CACATTGGCT	TGTACAGCGG	TGTTCCCAAG	TCGGTTCGGG	CCTTGGGGGT	360
GCAGCGCATT	TGCGGCACAG	GCTTCGAACT	GCTTCGGCAG	GCCGGCGAGC	AGATTTCCCA	420
AGGCGCTGAT	CACGTGCTGT	GTGTCGCGGC	AGAGTCCATG	TCGCGTAACC	CCATCGCGTC	480
GTATACACAC	CGGGGCGGGT	TCCGCCTCGG	TGCGCCCGTT	GAGTTCAAGG	ATTTTTTGTC	540
GGAGGCATTG	TTTGATCCTG	CTCCAGGACT	CGACATGATC	GCTACCGCAG	AAAACCTGGG	600
GGAGAGGCGG	TTTGCCTATT	GGGCGCATGC	ATAAAAACTG	TTGTAATTCA	TTAAGCATTC	660
TGCCGACATG	GAAGCCATCA	CAAACGGCAT	GATGAACCTG	AATCGCCAGC	GGCATCAGCA	720
CCTTGTCGCC	TTGCGTATAA	TATTTGCCCA	TGGACGCACA	CCGTGGAAAC	GGATGAAGGC	780
ACGAACCCAG	TTGACATAAG	CCTGTTTCGGT	TCGTAAACTG	TAATGCAAGT	AGCGTATGCG	840
CTCACGCAAC	TGGTCCAGAA	CCTTGACCGA	ACGCAGCGGT	GGTAACGGCG	CAGTGGCGGT	900
TTTCATGGCT	TGTTATGACT	GTTTTTTTGT	ACAGTCTATG	CCTCGGGCAT	CCAAGCAGCA	960
AGCGCGTTAC	GCCGTGGGTC	GATGTTTGAT	GTTATGGAGC	AGCAACGATG	TTACGCAGCA	1020
GCAACGATGT	TACGCAGCAG	GGCAGTCGCC	CTAAAACAAA	GTTAGGTGGC	TCAAGTATGG	1080
GCATCATTCT	CACATGTAGG	CTCGGCCCTG	ACCAAGTCAA	ATCCATGCGG	GCTGCTCTTG	1140
ATCTTTTCGG	TCGTGAGTTC	GGAGACGTAG	CCACCTACTC	CCAACATCAG	CCGGACTCCG	1200
ATTACCTCGG	GAACCTGCTC	CGTAGTAAGA	CATTCATCGC	GCTTGCTGCC	TTCGACCAAG	1260
AAGCGGTTGT	TGGCGCTCTC	GCGGCTTACG	TTCTGCCCAG	GTTTGAGCAG	CCGCGTAGTG	1320
AGATCTATAT	CTATGATCTC	GCAGTCTCCG	GCGAGCACCG	GAGGCAGGGC	ATTGCCACCG	1380
CGCTCATCAA	TCTCCTCAAG	CATGAGGCCA	ACGCGCTTGG	TGCTTATGTG	ATCTACGTGC	1440
AAGCAGATTA	CGGTGACGAT	CCCGCAGTGG	CTCTCTATAC	AAAGTTGGGC	ATACGGGAAG	1500
AAGTGATGCA	CTTTGATATC	GACCCAAGTA	CCGCCACCTA	ACAATTCTGT	CAAGCCGAGA	1560
TCGGCTTCCC	ATTGAGGGCG	CAAGAGGAGA	AATGGATTGA	CCAAGAGATC	GTGGCTGTTA	1620
CGGATGAACA	GTTTCGATTTA	GAGGGCTACA	ACAGTCGAGC	AATTGAACTG	CCTCGGAAGG	1680
CAAAATTGTT	GATCGTGACA	GTCATCCGCG	GCCTAGCAGT	CTTTGAAGCC	CTTTCCCGAT	1740
TGAAGCCTGT	TCATTCTGGC	GGGGTGCGAG	CTGCGGGCAA	CAGCTGTGCC	GTAGTGGACG	1800
GCGCCGCGGC	GGCTTTGGTG	GCTCGAGAGT	CGTCTGCGAC	ACAGCCGGTC	TTGGCTAGGA	1860
TACTGGCTAC	CTCCGTAGTC	GGGATCGAGC	CCGAGCATAT	GGGGCTCGGC	CCTGCGCCCG	1920
CGATTGCGCT	GCTGCTTGCG	CGTAGTGATC	TTAGTTTGAG	GGATATCGAC	CTCTTTGAGA	1980
TAAACGAGGC	GCAGGCCGCC	CAAGTTCTAG	CGGTACAGCA	TGAATTGGGT	ATTGAGCACT	2040
CAAAACTTAA	TATTTGGGGC	GGGGCCATTG	CACTTGGACA	CCCGCTTGCC	GCGACCGGAT	2100
TGCGTCTCTG	CATGACCCTC	GCTCACCAAT	TGCAAGCTAA	TAACTTTCGA	TATGGAATTG	2160

CCTCGGCATG	CATTGGTGGG	GGACAGGGGA	TGGCGGTTCT	TTTAGAGAAT	CCCCACTTCG	2220
GTTCGTCCTC	TGCACGAAGT	TCGATGATTA	ACAGAGTTGA	CCACTATCCA	CTGAGCTAAC	2280
GGGCATCTCC	TTTGTTGCTT	TGAGGTGGCG	CACGAAGGAG	GGCTCGAAAA	TCTCTGCTAA	2340
AAACAAGAAG	AAGGAACAGG	GAACATGATT	AGTTTCGCTC	GTATGGCAGA	AAGTTTAGGA	2400
GTCCAGGCTA	AACTTGCCCT	TGCCTTCGCA	CTCGTATTAT	GTGTCGGGCT	GATTGTTACC	2460
GGCACGGGTT	TCTACAGTGT	ACATACCTTG	TCAGGGTTGG	TGGGAATTC		2509

Sequenz 18

GAATTCCGCG	GTCGGCGAAA	GTTGATGCGC	TGTATCGTGG	TGAAGATCAA	TCCATGCTGC	60
GTGACGAGGC	CACACTGTGA	GTTGGTCAGG	GGGGGCTTAC	TCGGCGTTTT	CCGACACTGC	120
GTTGGTTGCG	GCAGTGCGCA	CCCCCTGGAT	TGATTGCGGG	GGTGCCCTGT	CGCTGGTGTC	180
GCCTATCGAC	TTAGGGGTAA	AGGTCGCTCG	CGAAGTTCTG	ATGCGTGCGT	CGCTTGAACC	240
ACAAATGGTC	GATAGCGTAC	TCGCAGGCTC	TATGGCTCAA	GCAAGCTTTG	ATGCTTACCT	300
GCTCCCGCGG	CACATTGGCT	TGTACAGCGG	TGTTCCCAAG	TCGGTTCCGG	CCTTGGGGGT	360
GCAGCGCATT	TGCGGCACAG	GCTTCGAACT	GCTTCGGCAG	GCCGGCGAGC	AGATTTCCCA	420
AGGCGCTGAT	CACGTGCTGT	GTGTCGCGGC	AGAGTCCATG	TCGCGTAACC	CCATCGCGTC	480
GTATACACAC	CGGGGCGGGT	TCCGCCCTCG	TGCGCCCGTT	GAGTTCAAGG	ATTTTTTGTG	540
GGAGGCATTG	TTTGATCCTG	CTCCAGGACT	CGACATGATC	GCTACCGCAG	AAAACCTGGC	600
GCGCATTGAG	GGCGCAAGAG	GAGAAATGGA	TTGACCAAGA	GATCGTGGCT	GTTACGGATG	660
AACAGTTCGA	TTTAGAGGGC	TACAACAGTC	GAGCAATTGA	ACTGCCTCGG	AAGGCAAAAT	720
TCTTCATCCT	CACACTGATC	CGGGGCGTAC	GAGTCTTTGA	ACCGCTTTCC	CGATTGATGC	780
CTGTTCATTC	TGGCGGGGTG	CAGACTGCGG	GCAACAGCTG	TGCCGTAGTG	GACGGCGCCG	840
CGGCGGCTTT	GGTGGCTCGA	GAGTCGTCTG	CGACACAGCC	GGTCTTGGCT	AGGATACTGG	900
CTACCTCCGT	AGTCGGGATC	GAGCCCAGC	ATATGGGGCT	CGGCCCTGCG	CCCGCGATTC	960
GCCTGCTGCT	TGCGCGTAGT	GATCTTAGTT	TGAGGGATAT	CGACCTCTTT	GAGATAAACG	1020
AGGCGCAGGC	CGCCCAAGTT	CTAGCGGTAC	AGCATGAATT	GGGTATTGAG	CACTCAAAAC	1080
TTAATATTTG	GGGCGGGGCC	ATTGCACTTG	GACACCCGCT	TGCCGCGACC	GGATTGCGTC	1140
TCTGCATGAC	CCTCGCTCAC	CAATTGCAAG	CTAATAACTT	TCGATATGGA	ATTGCCTCGG	1200
CATGCATTGG	TGGGGGACAG	GGGATGGCGG	TTCTTTTAGA	GAATCCCCAC	TTCGGTTCGT	1260
CCTCTGCACG	AAGTTCGATG	ATTAACAGAG	TTGACCACTA	TCCACTGAGC	TAACGGGCAT	1320
CTCCTTTGTT	GCTTTGAGGT	GGCGCACGAA	GGAGGGCTCG	AAAATCTCTG	CTAAAAACAA	1380
GAAGAAGGAA	CAGGGAACAT	GATTAGTTTC	GCTCGTATGG	CAGAAAGTTT	AGGAGTCCAG	1440
GCTAAACTTG	CCCTTGCTCT	CGCACTCGTA	TTATGTGTCTG	GGCTGATTGT	TACCGGCACG	1500
GGTTTCTACA	GTGTACATAC	CTTGTACAGG	TTGCTGGGAA	TTC		1543

Patentansprüche

1. Transformierter und/oder mutagenisierter ein- oder mehrzelliger Organismus, der dadurch gekennzeichnet ist, daß Enzyme des Eugenol- und/oder Ferulasäure-Katabolismus derart inaktiviert sind, daß eine Akkumulation der Intermediate Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure erfolgt.
2. Organismus nach Anspruch 1, dadurch gekennzeichnet, daß der Eugenol- und/oder Ferulasäure-Katabolismus durch Ω -Element-Insertion oder Einführen von Deletionen in entsprechende Gene verändert ist.
3. Organismus nach einem der Ansprüche 1 oder 2, dadurch gekennzeichnet, daß ein oder mehrere Gene, die für die Enzyme Coniferylalkohol-Dehydrogenasen, Coniferylaldehyd-Dehydrogenasen, Ferulasäure-CoA Synthetasen, Enoyl-CoA Hydratasen-Aldolasen, beta-Ketothiolasen, Vanillin-Dehydrogenasen oder Vanillinsäure-Demethylasen Enzyme kodieren, verändert und/oder inaktiviert sind.
4. Organismus nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß er einzellig, vorzugsweise ein Mikroorganismus oder eine pflanzliche oder eine tierische Zelle ist.
5. Organismus nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, daß er ein Bakterium, vorzugsweise eine *Pseudomonas*-Art ist.
6. Gen-Strukturen, bei denen die für die Enzyme Coniferylalkohol-Dehydrogenasen, Coniferylaldehyd-Dehydrogenasen, Ferulasäure-CoA Synthetasen, Enoyl-CoA Hydratasen-Aldolasen, beta-Ketothiolasen, Vanillin-Dehydrogenasen oder Vanillinsäure-Demethylasen oder zweier oder mehrerer dieser Enzyme kodierenden Nukleotidsequenzen verändert und/oder inaktiviert sind.

7. Gen-Strukturen mit den in Figur 1a bis 1r angegebenen Strukturen.
8. Gen-Strukturen mit den in Figur 2a bis 2r angegebenen Sequenzen.
- 5
9. Vektoren enthaltend wenigstens eine Gen-Struktur nach einem der Ansprüche 6 bis 8.

10. Transformierter Organismus nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß er wenigstens einen Vektor gemäß Anspruch 9 enthält.

11. Organismus nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß er wenigstens eine Gen-Struktur nach einem der Ansprüche 6 bis 8 an Stelle des jeweiligen intakten Gens im Genom integriert enthält.

12. Verfahren zur biotechnischen Herstellung von organischen Verbindungen, insbesondere von Alkoholen, Aldehyden und organischen Säuren, dadurch gekennzeichnet, daß ein Organismus nach einem der Ansprüche 1 bis 5 oder 10 bis 11 eingesetzt wird.

13. Verfahren zur Herstellung der Organismen nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß die Veränderung des Eugenol- und/oder Ferulasäure-Katabolismus mittels an sich bekannter mikrobiologischer Züchtungsmethoden erzielt wird.

14. Verfahren zur Herstellung eines Organismus nach einem der Ansprüche 1 bis 5 oder 10 bis 11, dadurch gekennzeichnet, daß die Veränderung des Eugenol- und/oder Ferulasäure-Katabolismus und/oder die Inaktivierung der entsprechenden Gene mittels gentechnischer Methoden erzielt wird.

15. Verwendung der Organismen nach einem der Ansprüche 1 bis 5 oder 10 bis 11 zur Herstellung von Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure.
 - 5 16. Verwendung von Gen-Strukturen nach einem der Ansprüche 6 bis 8 oder eines Vektors nach Anspruch 9 zur Herstellung transformierter und/oder mutagenisierter Organismen.
-

Konstruktion von Produktionsstämmen für die Herstellung von substituierten Phenolen durch gezielte Inaktivierungen von Genen des Eugenol- und Ferulasäure-Katabolismus

Zusammenfassung

Die vorliegende Erfindung betrifft einen transformierten und/oder mutagenisierten ein- oder mehrzelligen Organismus, der dadurch gekennzeichnet ist, daß Enzyme des Eugenol- und/oder Ferulasäure-Katabolismus derart inaktiviert sind, daß eine Akkumulation der Intermediate Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure erfolgt.

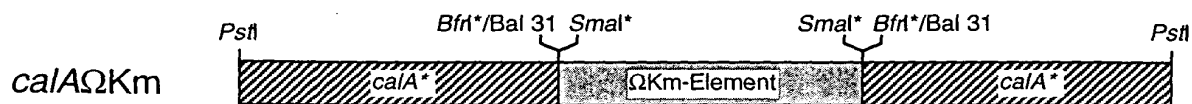


Fig. 1a

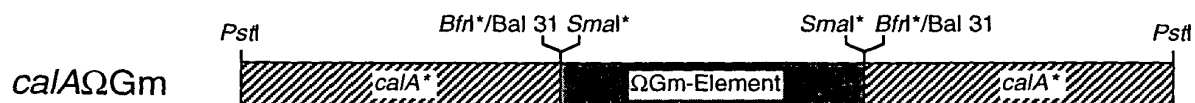


Fig. 1b

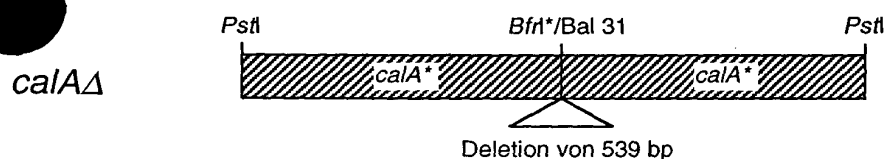


Fig. 1c

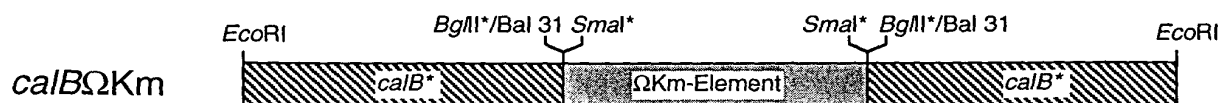


Fig. 1d

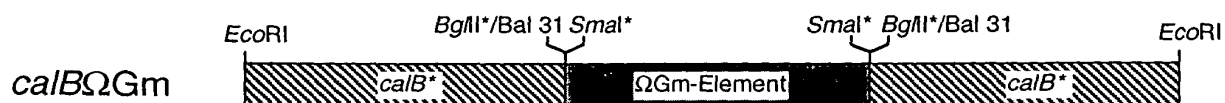


Fig. 1e

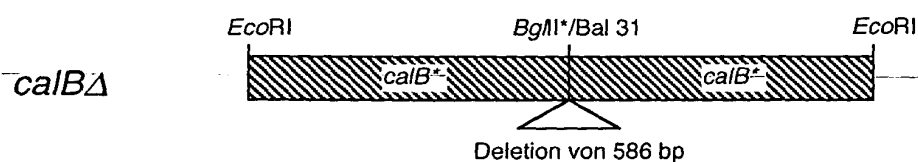


Fig. 1f

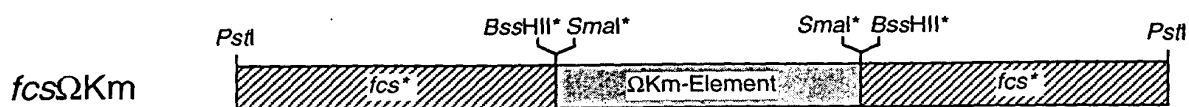


Fig. 1g



Fig. 1h

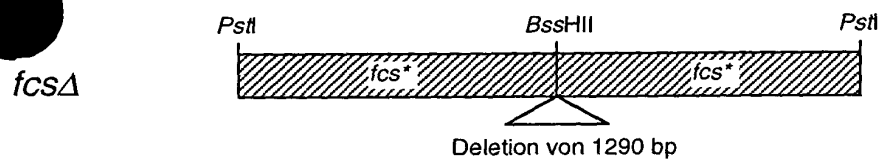


Fig. 1i

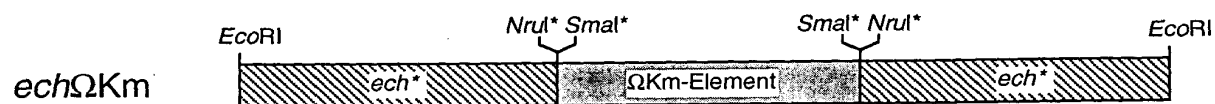


Fig. 1j

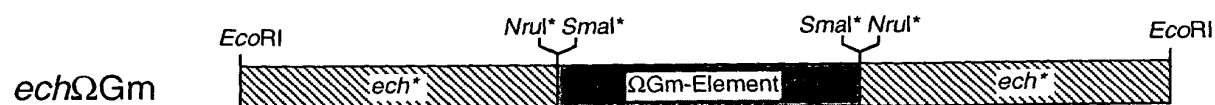


Fig. 1k

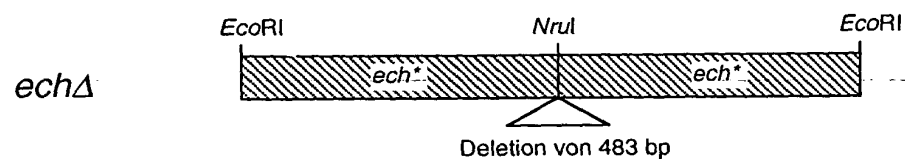


Fig. 1l

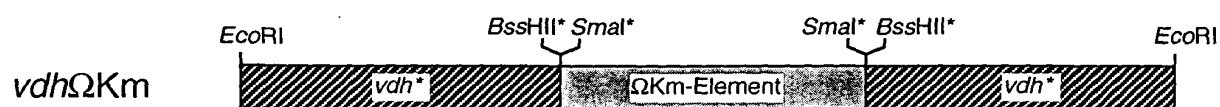


Fig. 1m



Fig. 1n



Deletion von 210 bp

Fig. 1o

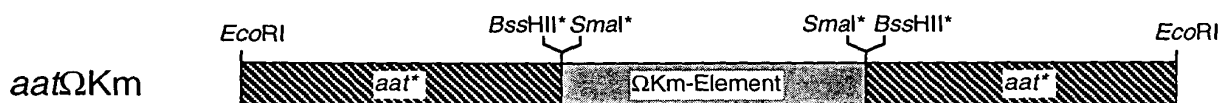


Fig. 1p

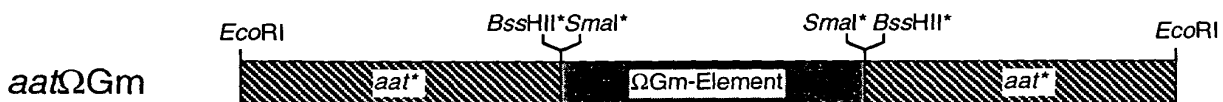


Fig. 1q



Deletion von 59 bp

Fig. 1r